U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANOGRAPHIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL CENTERS FOR ENVIRONMENTAL PREDICTION

original

OFFICE NOTE 419

THE USE OF COMPLEX QUALITY CONTROL FOR THE DETECTION AND CORRECTION OF ROUGH ERRORS IN RAWINSONDE HEIGHTS AND TEMPERATURES: A NEW ALGORITHM AT NCEP/EMC

William G. Collins
ENVIRONMENTAL MODELING CENTER

APRIL 1997

Abstract

Complex quality control (cqc) was a new concept when introduced to NCEP by Lev Gandin in the late 1980's. The basic idea is that no decisions are to be made on data quality until the results from all checks, in numerical form, are available. In addition, data are to be corrected when possible. Since Lev's arrival at NCEP, the author has worked with him to introduce cqc for the quality control of rawinsonde heights and temperatures and Doppler profiler winds. The principles were also used in the quality control of many other data types.

This note describes the code that resulted from rethinking the strategy for rawinsonde heights and temperatures. The mandatory and significant levels are now treated together, rather than first all the mandatory levels, and then the significant levels. The data are considered, one level at a time, from the ground upward. There is also a greater opportunity for complex errors to be corrected since, whenever an error is corrected, then all residuals are recalculated and computations begin anew from the lowest level. Only after all possibly correctable errors are considered are observational errors looked for.

The note begins with a general consideration of errors, then turns to methods of error detection and correction. The detail becomes higher as the code specifics are described. Sample output from the code is described, examples are given, and statistical results of use of the code are presented.

1. Introduction

Errors in meteorological data come from several sources. There are instrument and processing errors, such as rounding which are largely random, unbiased and small in magnitude. These errors are unavoidable, and are also unimportant for almost all uses of the data. There are so-called errors of representativeness. These are not actually errors at all but the sampling by the instrument of a local value of the variable whose scale is not representative of the environment for a particular use of the data. For instance, the temperature within the updraft of a thunderstorm would not be representative of the environment for use of the data by a global forecast model assimilation. The third category of errors will be called rough errors. It is the quality control (qc) of these errors, to be described in some detail in the next paragraph, that are the subject of this note, especially those in the height and temperature of rawinsonde soundings.

Rough errors, as the name implies, are generally moderately to extremely large. They are not random, and bias is not a suitable measure for them. There are two main classes of rough errors, stemming from their origin: those made by human action, and those due to observation instrument failure of some kind.

The rough errors due to human action may be due to the incorrect writing of a number or incorrect coding of a value. Errors of this nature are referred to as communication errors. Rough errors may also be due to a computation error, e.g. in computing heights. Both communication and computation errors may usually be corrected.

Rough observation errors are generally due to gross instrument failure or miss-calibration. However, observation errors of moderate magnitude may sometimes be misidentified since the measures to evaluate them are not perfect and also because they can be confused with errors of representativeness.

The WMO code for the transmission of upper-air data requires that both geopotential height and temperature, along with dew-point depression, be sent in the message. The inclusion of heights and temperatures/dew-point temperatures is redundant since the heights are hydrostatically calculated at each station from the temperatures and dew-point temperatures at known pressures. Rough errors of communication or computation type may therefore be diagnosed and an error correction suggested by the hydrostatic inconsistencies between the reported heights and temperatures.

The objective of complex quality control for rawinsonde heights and temperatures (cqcht) is to detect rough errors--communication, computation and observation errors--and to correct as many communication and computation errors as possible. It is sometimes impossible to distinguish errors of representativeness from observation errors. However, it is the intention **not** to identify errors of representativeness. (See Gandin, 1988 for a discussion of errors.)

Many of the possible errors in rawinsonde observations are illustrated in Fig. 1. There are errors than can occur in instrument manufacture or calibration and instrument failure. Many types of errors can take place at the observing cite, whether the data is processed manually or automatically. And there can be errors en route to and at national processing centers. In the figure a superscript **c** indicates a "communication" error and **o** indicates an "observation" error.

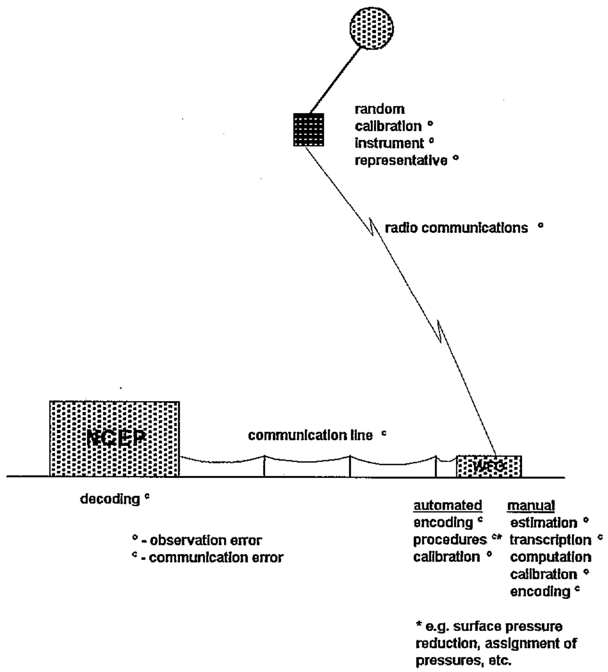

Fig. 1 Errors associated with rawinsonde observations and processing.

Section 2 gives the general principles of error detection. It is seen that it is the properties of meteorological data itself that allows errors to be located. The next section, Section 3, goes on the tell the various checks that are used in error detection. More specific ideas on error detection principles of cqc are given in Section 4. Section 5 goes into the methods of error correction and complex quality control. Section 6 goes into more detail with the design characteristics of the algorithm that makes the actual quality and correction decisions, the so-called Decision Making Algorithm (DMA). Section 7 outlines the low level routines that make the actual error corrections. The various outputs that are produced are described in Section 8. Examples of various errors are given in Section 9, and statistics of error detection and correction are shown in Section 10. The general reader may wish to skip Sections 6-8, since they progressively go into more detail. The user of the results of cqcht96 will especially want to look at Section 8.

2. Detection of rough errors

The detection of rough errors in rawinsonde heights and temperatures requires skills in several disciplines, including meteorology, mathematics, and psychology. For this reason, it may be likened to the work of a criminologist. As the story of the methods of error detection and correction unfolds, it is seen how knowledge from each of the disciplines is used.

Meteorological fields are, by and large, smooth, except in the vicinity of fronts, near convective systems, and near rapid surface property changes. The degree of this smoothness may be measured by correlating a variable, or its departure from climatology or a forecast, at a position with its value at other positions separated horizontally, vertically or temporally. The variation of such correlations is important in the design of data assimilation methods, and also for cqcht. It is only because the spatial scales of such correlations are much larger than the resolution of assimilation models that an assimilation is at all possible. Without sufficiently large correlation scales for meteorological variables, several tools used in the qc of data would also be ineffective.

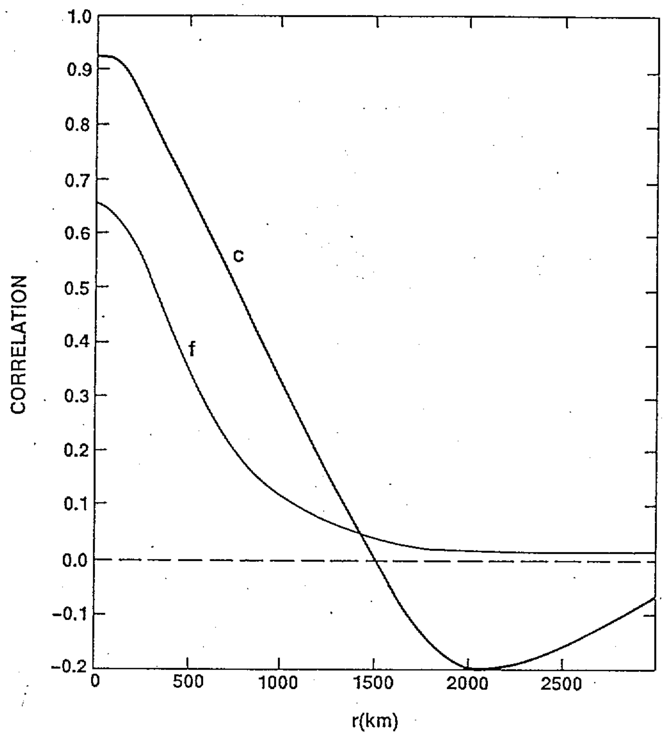

Fig. 2 Observed-minus-background correlation for the 500 hPa geopotential as a function of distance between stations. Curve c is for climatological background and curve f is for a forecast background. (Copied *from Atmospheric Data Analysis*, Daley, p. 114.)

Fig. 2 shows the variation with distance of the deviation of the 500 hPa geopotential from both climatology and a forecast. The correlations fall by half in about 500 km, thus implying that data at least that far away would be useful in checking a geopotential value. Similar correlations have been derived for vertical and temporal separations, with similar results. Naturally, the vertical correlations fall off rather rapidly, but rawinsondes have good vertical resolution, so the falloff to the nearest levels is usually not large.

A vertical temperature profile will generally contain lapse rates that are stable or at most slightly unstable. This may be used to diagnose temperatures that are suspect and to check "corrected" temperatures.

The redundancy in the reported heights and temperatures and the use of the hydrostatic approximation to calculate the geopotential heights, in the first place, leads to a powerful method for the diagnosis of errors. Hydrostatic inconsistencies between the reported geopotential heights and temperatures may be used to determine which datum is in error and to suggest a correction. The **lack** of hydrostatic inconsistency, likewise, may be used to imply that any errors are not communication or computation errors, but rather observation errors.

In modern forecast models, the forecast heights and temperatures are highly accurate and therefore any large deviation of data from the forecast makes them suspect, especially if the difference from the first forecast changes horizontally or vertically. The difference of the data from the forecast will be used in several of the checks to diagnose errors, sometimes to suggest corrections, and to validate any corrections made. The next section will give more detail on the various checks. Previous results of monitoring for rough errors in reports was reported in Gandin, et al, 1993

3. The checks

The various checks use the properties of meteorological fields and the redundancy of heights and temperatures, along with the hydrostatic approximation to make independent assessment of data quality. The checks are: increment, horizontal statistical, vertical statistical, lapse rate, hydrostatic and baseline. The remainder of this section will define the check results. Earlier versions of cqc for rawinsonde heights and temperatures included most of the same checks (see Collins and Gandin, 1990; Collins and Gandin, 1992).

a. Increment

Recognizing the value of a 6-hour forecast in identifying errors, the increment (or observed increment) is calculated, defined as the difference between a reported value and its forecast value, interpolated both horizontally and vertically to the data location. Within NCEP's operations the data is provided to cqcht96 in a locally defined BUFR format called prepbufr. In prepbufr, the guess (6-hour forecast) values,

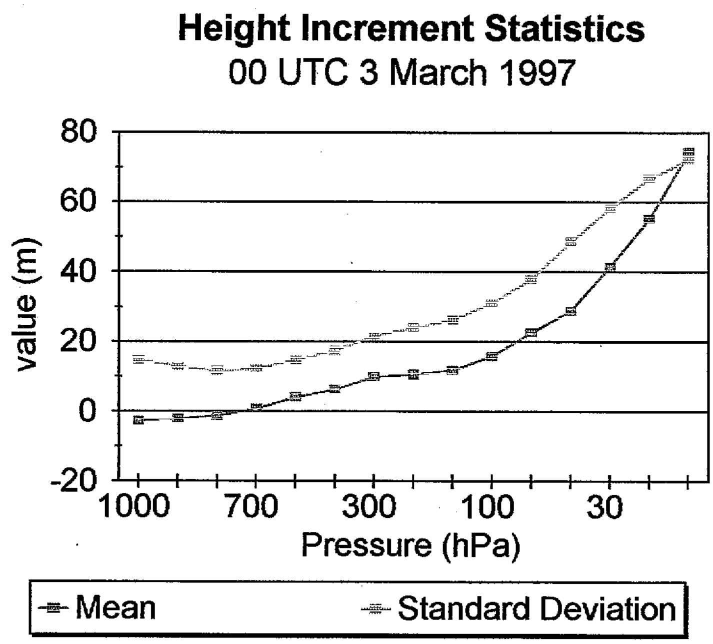

Fig. 3 Height increment mean and standard deviation for all stations for a single time: 00 UTC 3 March 1997.

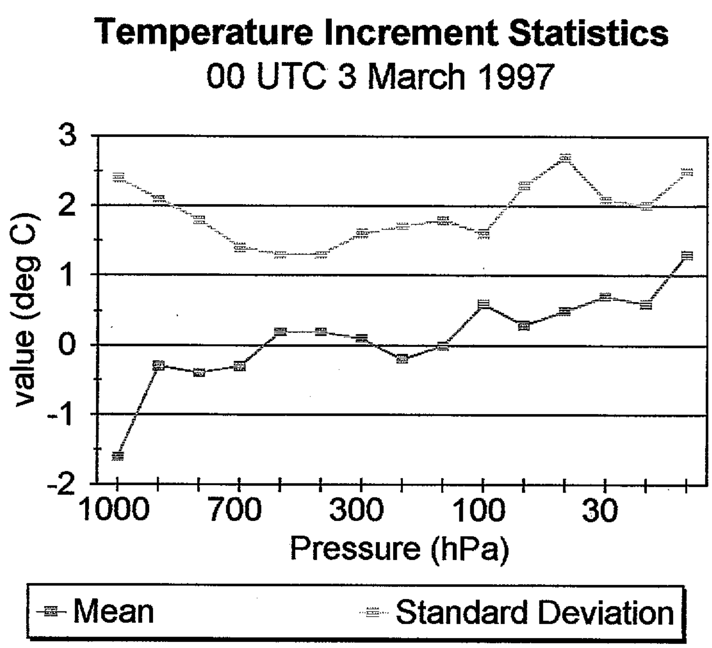

Fig. 4 Temperature increment mean and standard deviation for all stations for a single time: 00 UTC 3 March 1997.

interpolated to data locations, are already part of the file when used by cqcht96. Thus, the calculation of the increment is just

$$I = O - G$$

for each height and temperature, where I is the observed increment, O is the observed value, and G is the 6-hour forecast value. Figs. 3 and 4 shows sample statistics for the height and temperature increments. They show the mean and standard deviations for all observations for 00 UTC 3 March 1997, stratified by pressure.

For errors at the surface, it is useful to define the surface pressure increment, the difference between the reported surface pressure and the 6-hour forecast surface pressure, vertically adjusted from the model terrain to the station elevation.

b. Horizontal statistical residual

The horizontal statistical check performs horizontal optimal interpolation. A value of increment is interpolated from at most four surrounding values, the closest one from each quadrant (less than 1000 km distance), to the observed increment location. The horizontal residual is defined as the difference between the observed increment and the interpolated increment.

The horizontal residual is calculated as

$$S_l^h = I_l - \sum_{i=1}^{4} w_i I_i$$

where S_l^h is the horizontal residual (at a point l), I_l is the observed increment, and w_i are the weights, determined from

$$\begin{vmatrix} 1+\varepsilon & r_{12} & r_{13} & r_{14} \\ r_{21} & 1+\varepsilon & r_{23} & r_{24} \\ r_{31} & r_{32} & 1+\varepsilon & r_{34} \\ r_{41} & r_{42} & r_{43} & 1+\varepsilon \end{vmatrix} \begin{vmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{vmatrix} = \begin{vmatrix} r_{01} \\ r_{02} \\ r_{03} \\ r_{04} \end{vmatrix}$$

where the unknown weights are w, ε is the ratio of 6-hour observation error variance to forecast error variance, given the value 0.5, and r_{ij} is the correlation between the increments at points i and j. The observation point is denoted by the subscript 0. The correlations are modeled with a squared exponential that depends only upon distance. It is shown in Fig. 5 and is given by

$$r_{ij} = \exp(-kd_{ij}^2).$$

The constant, k, has the value $3.5\mathrm{x}10^{-6}\ \mathrm{m}^{-2}$. The equation for the weights is solved by a standard matrix method.

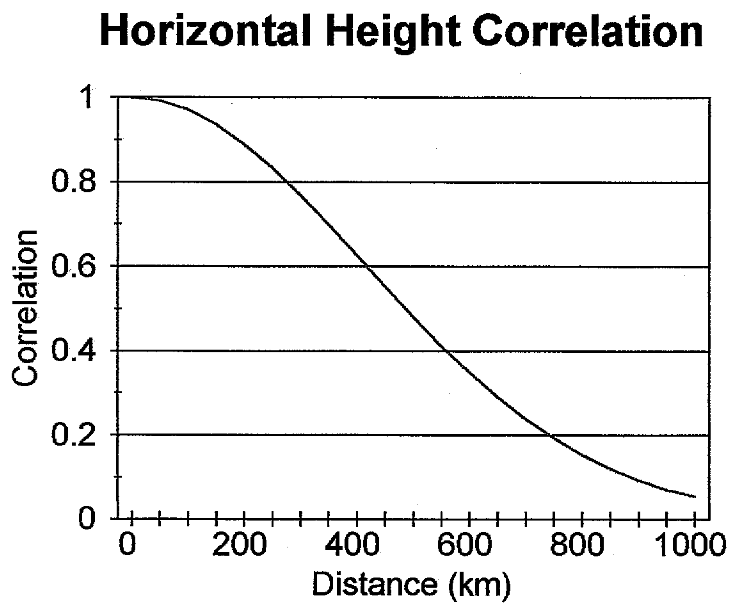

Fig. 5 Horizontal height auto-correlation as a function of distance assumed by horizontal check.

c. Vertical statistical residual

The vertical residual is the difference between the observed increment and the increment interpolated vertically from the nearest data points for the same station, one above and one below. Thus, the vertical residual is given by

$$S_l^v = I_l - w_{l-1}I_{l-1} - w_{l+1}I_{l+1}$$

where l-1 is the first level below and l+1 is the first level above the data level, l. The weights are determined to give minimal rms. error. For a formulation of the problem and solution, see Thiebaux and Pedder, 1987, chapter 4. The weights are given by

$$w_{l+1} = \frac{(1+\gamma)r_{l,l+1} - r_{l,l-1}r_{l-1,l+1}}{(1+\gamma)^2 - r_{l-1,l+1}^{\ 2}} \text{ and } w_{l-1} = \frac{(1+\gamma)r_{l,l-1} - r_{l,l+1}r_{l-1,l+1}}{(1+\gamma)^2 - r_{l-1,l+1}^{\ 2}}$$

where $r_{i,j}$ is the vertical correlation of increments between level i and j and $\gamma = 0.5$ is the assumed ratio of the observation to 6-hour forecast error variance. The vertical correlation model used is

$$r_{l1,l2} = \frac{1}{1 + c_a \left| \ln\left(\frac{p_{l1}}{p_{l2}}\right) \right|^{1.2}}.$$

The value of c_a is 1.1 for height and 8.0 for temperature. Fig. 6 shows the correlations as a function of the ratio of the two pressures. The height correlation falls off much more slowly because of the large vertical correlation of the height increment through the hydrostatic approximation.

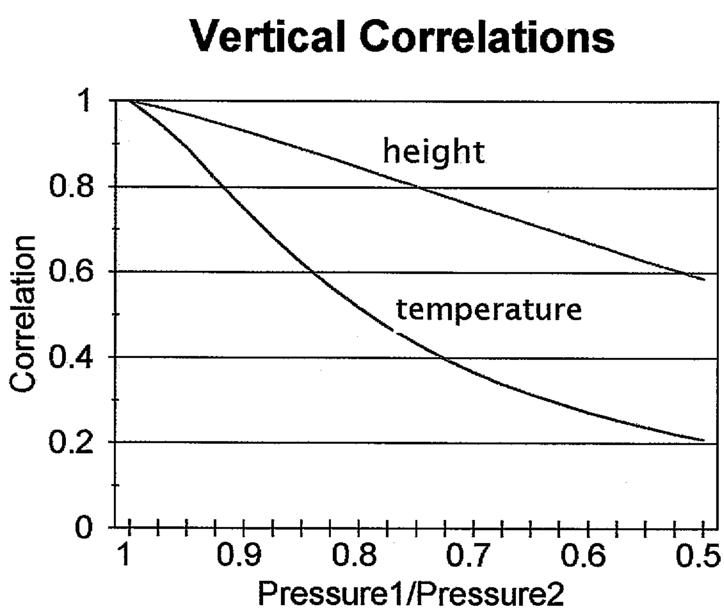

Fig. 6 Assumed vertical variation of auto-correlation of height and temperature as a function of the ratio of the two pressures. These variations are used by the vertical checks.

d. Hydrostatic residual

As stated previously, the complex of hydrostatic residuals is necessary for determination of error types as well as providing the information needed for accurate error correction. It is noted that a hydrostatic residual is only available between mandatory levels where both height and temperature are available. The hydrostatic residual is the difference between the thickness computed by the heights, and the thickness computed from the temperatures. When there are no communication or communication errors, the hydrostatic residuals are small. There are several variants of hydrostatic residual that are possible depending upon whether:

a. only mandatory level information is used, or significant level temperatures are also used,
b. moisture data (dew-point temperature depression) are used.

The problem with using moisture data is that it is almost impossible to quality control. However, in cqcht96 it is used if its use does not greatly modify the value of the hydrostatic residual, compared to its value without its use. In the following equations, it is assumed that the temperature is the virtual temperature, and two forms will be given. In the first, only the layer bounding level mandatory heights and temperatures are used. In the second, any intermediate significant level temperatures are also used.

The hydrostatic residual for a layer between two mandatory levels, *l1* and *l2*, each containing a height and temperature, and using no intervening significant levels, is

$$S^{m}_{l1,l2} = z_{l2} - z_{l1} - A_{l1,l2} - B_{l1,l2}(T_{l1} + T_{l2})$$

where T is the (virtual) temperature in Celsius and z the geopotential height. The coefficients A and B are given by

$$A_{l1,l2} = \frac{RT_o}{g}\ln\left(\frac{p_{l1}}{p_{l2}}\right) \text{ and } B_{l1,l2} = \frac{R}{2g}\ln\left(\frac{p_{l1}}{p_{l2}}\right)$$

where T_o = 273.15 K, R is the gas constant for dry air and g is the acceleration of gravity.

When any intervening significant temperatures are also used, the hydrostatic residual is

$$S^{s}_{l1,l2} = z_{l2} - z_{l1} - A_{l1,l2} - \sum_{i=l1}^{l2-1} B_{i,i+1}\left(T_i + T_{i+1}\right)$$

where $A_{l1,l2}$ has the same value as before and

$$B_{i,i+1} = \frac{R}{2g}\ln\left(\frac{p_i}{p_{i+1}}\right).$$

In the absence of error in any of the quantities used in computing it, the hydrostatic residual is small, but not necessarily zero since the full set of temperatures used at each reporting station to compute the heights is generally not available.

Sometimes it is useful to scale the hydrostatic residual so that it has the units of temperature. Such scaling will be seen in the examples. It is done by dividing the hydrostatic residual by *B*:

$$X_{l1,l2} = S_{l1,l2} / B_{l1,l2}$$

e. Baseline residual

The baseline residual is a measure of the mismatch between the lower heights and temperatures of the upper-air report and the station elevation. It is computed by making a hydrostatic computation downward from the first complete (i.e. level with non-missing height and temperature) mandatory level above the surface to the reported surface pressure. The baseline residual, S^b, is the difference between the station elevation, given by the report, and the hydrostatically determined height at the surface pressure. It is

$$S^b = z_1 - z_s + \sum_i \left[\frac{RT_0}{g} + \frac{R}{2g}\left(T_i + T_{i+1}\right)\ln\left(\frac{p_{i+1}}{p_i}\right)\right]$$

where the sum is over the layers as stated above and z_1 is the height of the first complete mandatory level above the surface. The (virtual) temperature is in Celsius. The baseline residual has units of height.

It is sometimes convenient to consider the baseline data mismatch from a different point of view. One may ask what the pressure inconsistency is between the reported surface pressure and the pressure obtained when working down hydrostatically from the first complete mandatory level to the reported station elevation. By this computation the surface pressure residual is obtained, given by

$$S^p = p_s\left\{1-\exp\left[\frac{2g}{R}\left(\frac{z_0 - z_s}{2T_0 + T_s + T_{s+}}\right)\right]\right\}$$

with

$$z_0 \equiv z_1 + \sum_i\left[\frac{RT_o}{g} + \frac{R}{2g}(T_i + T_{i+1})\ln\left(\frac{p_i}{p_{i+1}}\right)\right]$$

where s_+ refers to the first level above the surface. Other symbols have the same meaning as before, and the summation is over the same layers. Fig. 7 shows the arrangement of data for this check.

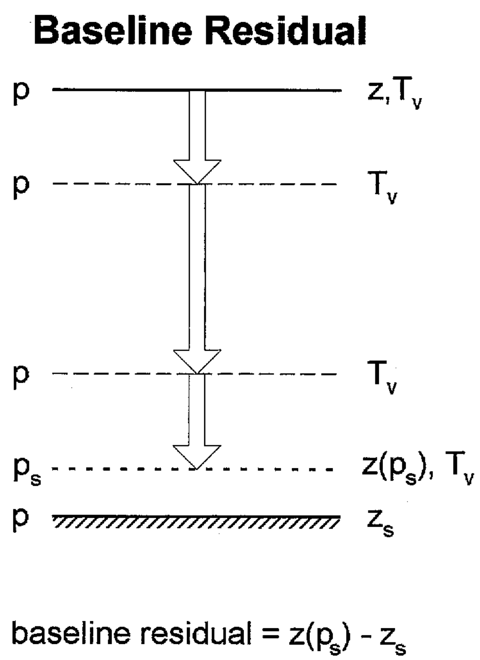

Fig. 7 Arrangement of variables for the baseline residual. The hydrostatic computation proceeds downward from the first complete mandatory level to the station elevation, using any intervening (virtual) temperatures.

f. Lapse rate check

The lapse rate is computed between each temperature and those immediately above or below. The lapse rate is placed in one of four classes: 1) absolutely stable, 2) conditionally stable (stable with respect to unsaturated air but unstable with respect to saturated air), 3) unstable, and 4) unstable with loose limits.

The division between absolutely stable and conditionally stable is defined by the dry-adiabatic lapse rate:

$$\gamma_d = \frac{1}{c_{pd}}$$

with

$$c_{pd} = 1004.5 \text{ J kg}^{-1} \text{ K}^{-1}$$

where c_{pd} is the specific heat at constant pressure for dry air. The lapse rate is conditionally stable if the lapse rate is between the dry-adiabatic lapse rate and the rate for the saturated adiabat.

The saturated adiabatic lapse rate is given by

$$\gamma_s = -\frac{dT_s}{dz} = \frac{g}{c_p} \frac{p + \dfrac{\varepsilon L e_s}{R_d T_K}}{p + \dfrac{\varepsilon}{c_p} \dfrac{L^2 e_s}{R_v T_K{}^2}}$$

with

$$e_s = 6.1078 \exp\left(\frac{17.269\, T_c}{T_c + 237.3}\right),$$

$$\varepsilon = 0.622$$

$$R_d = 287.05 \text{ J kg}^{-1} \text{ K}^{-1}$$

$$R_v = 461.5 \text{ J kg}^{-1} \text{ K}^{-1}$$

$$L = (2.501 - .00237\, T_c) \cdot 10^6 \text{ J kg}^{-1}$$

In these formulas, T_k is the temperature in Kelvin, while T_c is the temperature in Celsius. ($T_c = T_K - 273.15$)

The lapse rate is defined as *-dT/dz*. It is approximated between levels p_+ and p_-, with temperatures T_+ and T_-, by

$$\gamma = -\frac{dT}{dz} = -\frac{dT}{dp}\frac{dp}{dz} = \frac{g}{R}\frac{d\ln T}{d\ln p} \cong \frac{g}{R}\frac{\ln\left(\dfrac{T_+}{T_-}\right)}{\ln\left(\dfrac{p_+}{p_-}\right)},$$

and the lapse rate with loose limits is defined as

$$\gamma_{loose} \equiv \frac{g}{R}\frac{\ln\left(\dfrac{T_+ + 2.}{T_-}\right)}{\ln\left(\dfrac{p_+}{p_-}\right)}$$

In summary, the lapse rate classes are determined as follows:

4) If $\gamma_{loose} > g/c_p$, then the rate is absolutely unstable with loose limits,
3) Else if $\gamma > g/c_p$, then the rate is absolutely unstable,
2) Else if $\gamma > \gamma s$ and $\gamma < g/c_p$, then the rate is conditionally stable, and
1) Otherwise, the rate is absolutely stable.

The lapse rate only gives a limiting condition and therefore is not a very powerful check. It is used in the preliminary evaluation of temperatures to eliminate bad data from the horizontal check. And the lapse rate check is also used in the search for observation errors.

4. Detection of Errors—Complex Quality Control

Some general principles must be emphasized for cqc. All of the residuals are calculated before any decisions are made, and the cqc will use the agreement of the values of the various residuals, including increment, in making its decisions.

It is not only the residuals at the suspected error location that are important, but also the pattern of the residuals within the vicinity. This is true since, of all the residuals, only the increment does not involve data at other locations. The vertical, hydrostatic, lapse rate, and baseline checks all use data separated vertically, while the horizontal check uses data separated horizontally.

Put in another way, an error in height, say, will affect the increment at the data location, the hydrostatic residuals for the layers above and below (perhaps also the baseline residual), the vertical residuals for the data levels above and below, and if not properly eliminated from use, the horizontal residuals at neighboring points at the same level. A temperature error would, in addition, affect the lapse rates for the layers above and below.

Some of the dependency of the check results is illustrated in Figs. 8-10. Fig. 8 shows the effects upon the residuals for a temperature or height communication error at a mandatory level. The error in the data is assumed to be positive, indicated by "+". Note that the vertical check for temperature uses the nearest temperature above or below, whether they be at mandatory or significant levels. The horizontal check is only performed at mandatory levels. Two neighboring points are shown in the figure. The "-" at several of the data points is meant to indicate that the error would have an impact of negative sense upon checks at these points. For a positive temperature error, both hydrostatic residuals are positive, indicated by "+". If the error were in height, then the lower hydrostatic residual would be positive, and the upper hydrostatic residual would be negative.

Residuals at Mandatory Levels

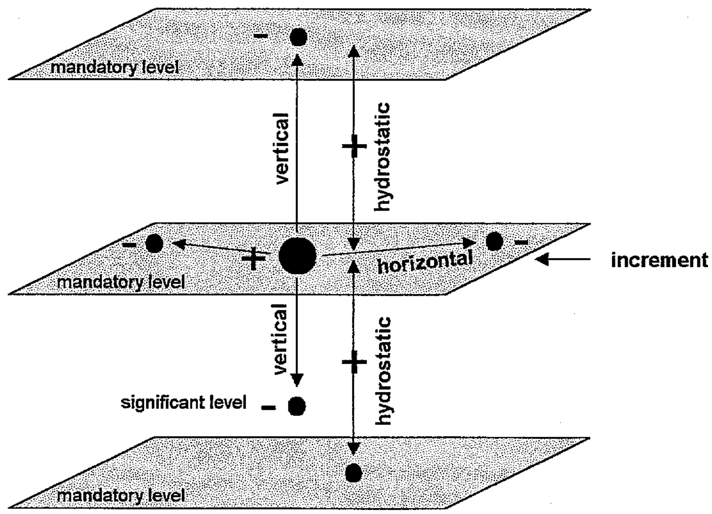

Fig. 8 Arrangement of variables for check at mandatory level. A positive height error is assumed. For a positive temperature error, the lower hydrostatic residual would be positive and the upper on negative.

Fig. 9 shows the residuals associated with a temperature communication error at a significant level. Again, the vertical check uses the closest temperatures above and below. The "-" at levels used in the vertical check indicates a negative influence upon vertical checks of the data at these levels. The lapse rate checks use the same neighboring levels. No horizontal check is available at significant levels. While hydrostatic checks cannot be made above and below this significant level, a hydrostatic check can be made that includes the level. This check residual will be affected by any communication error at this significant level and thus gives evidence of such an error.

Residuals at Significant Levels

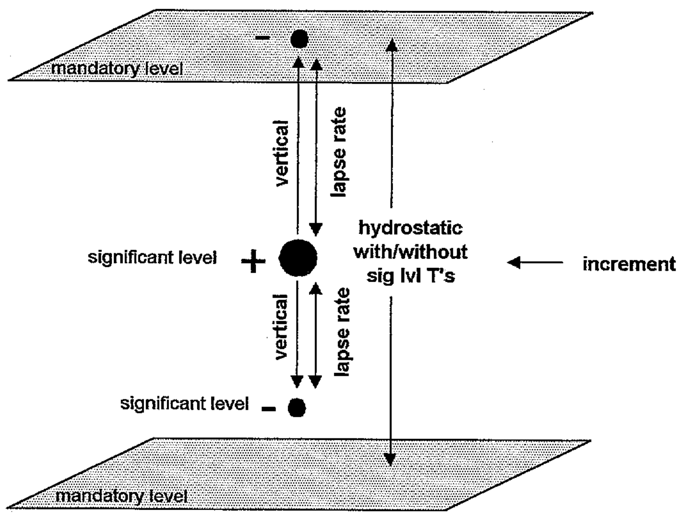

Fig. 9. Arrangement of variables for checks at significant levels.

Fig. 10 shows the effect of a height computation error upon various residuals. The vertical checks immediately above and below the data error level are affected by a computation error, but the magnitude of the residuals cannot be used to determine the magnitude of the error. Therefore, vertical checks will not be further considered for this kind of error. At levels below the error, hydrostatic residuals, increments and horizontal residuals are unaffected. For the layer where the error occurs, the hydrostatic residual has the magnitude of the error. At all levels above the error, the increments and horizontal residuals have (roughly) the magnitude of the error. A correction will modify all heights above the layer of the error by the same amount.

Residuals Due to Thickness Computation Error

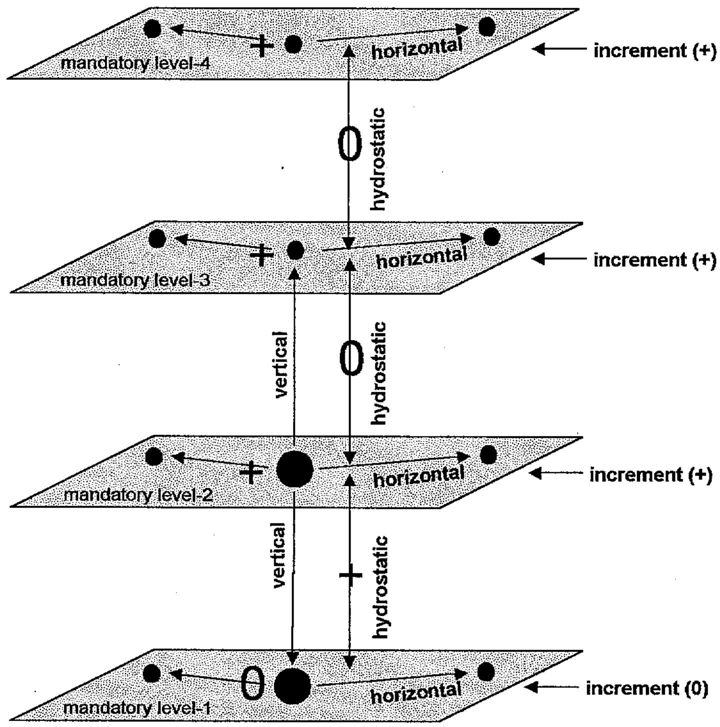

Fig. 10 Residuals when there is a computation error between levels 1 and 2. Note that all levels 2 and above are in error by the same amount.

The usefulness of each check for error determination may be measured by its mean and standard deviation. Long-term averages of these statistics are used in the decisions. Figs. 11 and 12 show the check means for height and temperature for a sample day. And Figs. 13 and 14 show the check standard deviations. The hydrostatic check has the smallest standard deviation, followed by the vertical, horizontal and increment checks. The temperature standard deviations vary little with pressure, while the height standard deviations grow for smaller pressure. The height check means also grow with elevation, while the temperature check means are mostly small.

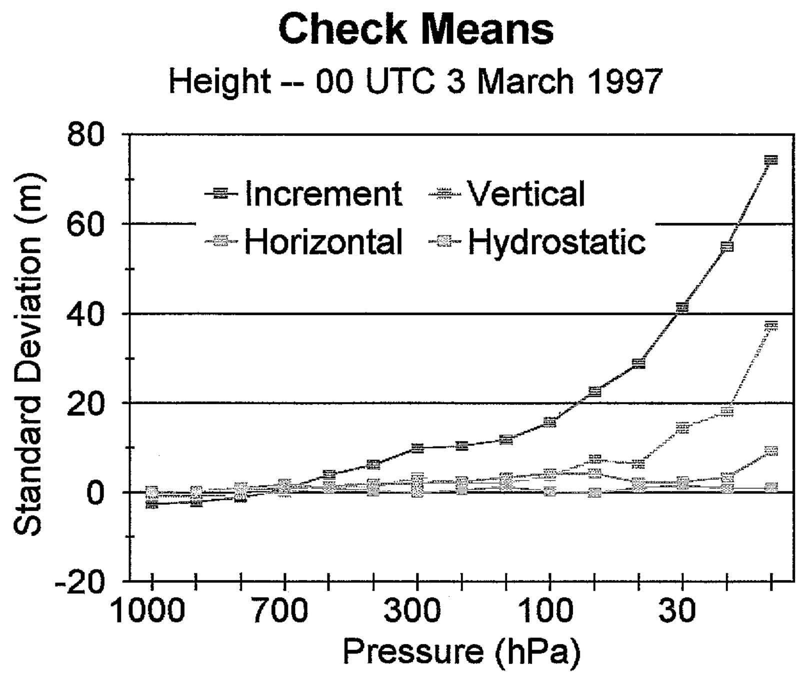

Fig. 11 Check means for height for a sample time: 00UTC 3 March 1997. Includes increment, horizontal, vertical and hydrostatic checks.

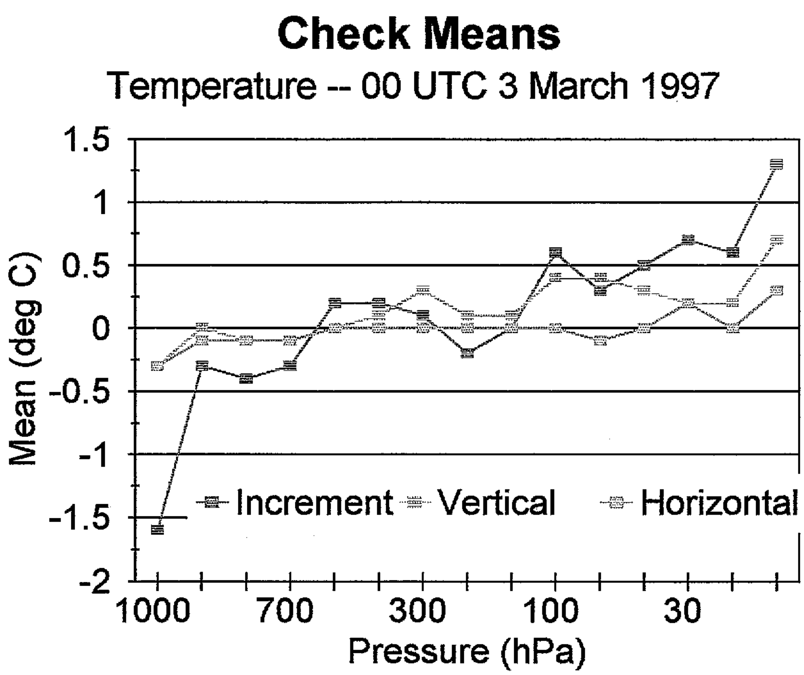

Fig. 12 Check means for temperature for a sample time: 00UTC 3 March 1997.

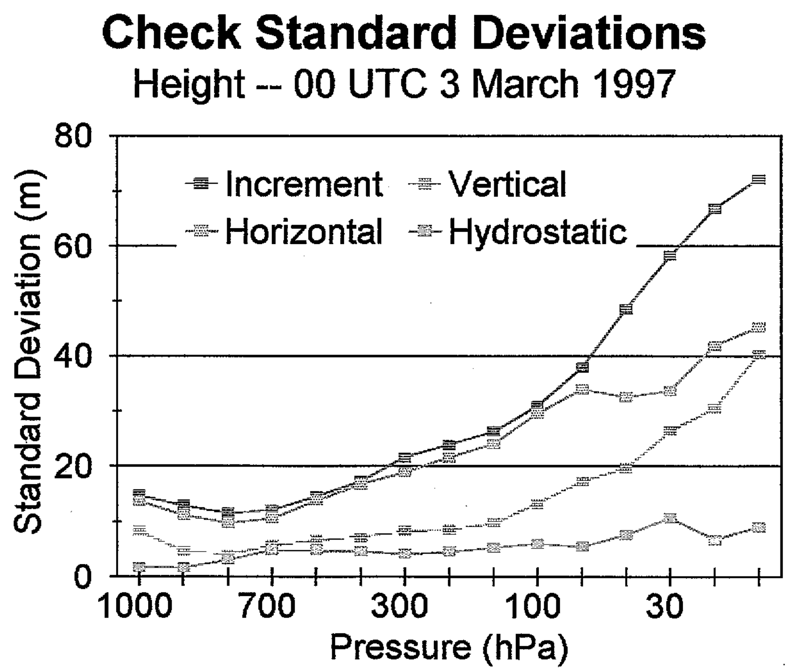

Fig. 13 Height check standard deviations for a sample time: 00UTC 3 March 1997.

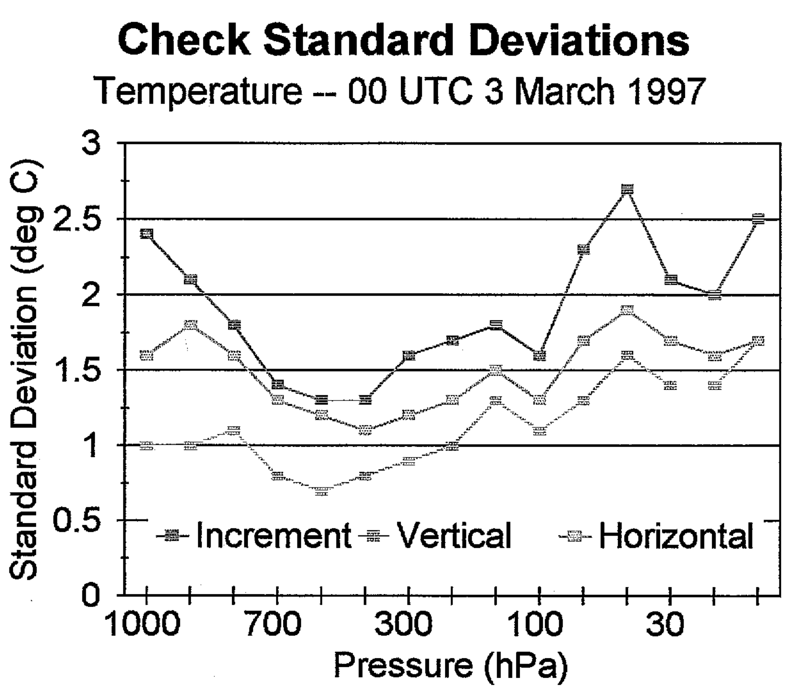

Fig. 13 Temperature check standard deviations for a sample time: 00UTC 3 March 1997.

5. Correction of Errors—Complex Quality Control

The previous section emphasized the detection of errors. This section will emphasize the principles used by cqc for their correction. The various check residuals not only give information on the location of an error, but also its magnitude. Therefore, it is essential that all residuals be available before decisions are made. More specific information of the design of the part of the algorithm that makes decisions, called the Decision Making Algorithm (DMA), will be given in the next sections. The DMA will use the complex of residuals in its decisions.

Most correctable errors may be diagnosed by the complex of hydrostatic residuals alone. However, other residuals are needed in some cases to distinguish which variables are in error and in all cases to confirm a suspected error. In any case, it is the hydrostatic residuals that distinguish between communication, computation, and observation errors. A computation error is signaled by a single large hydrostatic residual, communication errors by multiple large hydrostatic residuals, and observation errors by lack of large hydrostatic residuals.

The reported geopotential heights are calculated at each station from the observed temperatures, relative humidity and pressures. Accurate enough representation of the full set of observed values is present in the WMO coded message so that, in absence of error, the hydrostatic residuals are quite small. In fact, this is confirmed by the very small standard deviations for the hydrostatic check as already seen in Fig. 13. (The mean value of the hydrostatic check is practically 0., shown in Fig. 11.) Therefore, the redundancy in the reported heights and temperatures places the hydrostatic check in the unique position that its values do not merely give estimates of error magnitudes but rather gives highly accurate estimates of an error magnitudes.

The accurate estimate of the error magnitudes makes the attempt to recover an original value tenable. The attempt is made even more reasonable by the fact that most communication and computation errors come about by direct human action: writing down numbers or typing them incorrectly, miss-coding the data, making computation errors, etc. The most common errors are a sign error (to temperature), single digit error or interchange of digits. Cqcht96 specifically looks for corrections to errors of these kinds. Such corrections are called "simple" corrections. Inspection of the corrections that the code makes shows that most are simple.

A correction, to be valid, must fit surrounding data. Therefore, no correction is accepted unless all residuals are made smaller by its introduction. In most cases, the original correction magnitude is given by the complex of hydrostatic residuals. It may then be modified to make it simple. Next, the provisional correction is checked to see if all residuals are made smaller. If they are, then the correction is accepted. If they are not all smaller, then a new provisional correction is formed as a weighted average of the available residuals. This provisional correction is again possibly made simple and then tested for acceptability, being finally accepted or rejected. If, in the end, no correction is made, then the datum is flagged questionable or bad, depending upon the magnitudes of the residuals.

Any correction is allowed to influence later decisions, thus allowing in some cases for rather complicated corrections. The decisions are performed from the lowest to the highest level, and whenever a correction is made, all residuals are recalculated and the decision making begins anew from the bottom. In this way, several scans may be made, with one correction on each scan. Observation errors are considered only on a final last scan.

6. Design Characteristics of the Decision Making Algorithm

Some of the characteristics of the DMA have already been outlined. This section will go into considerably more detail.

The quality control is performed from the lowest level upward, including both significant and mandatory levels. When a correction is made, then the residuals are recalculated and a new scan through the data is begun from the bottom.

An early concern in the design of the DMA was to make this cqcht96 as generally applicable as possible. On way to do that would be to make it, so far as possible, station independent. Now, the only check that requires information from other stations is the horizontal check. Further, it was seen in Figs. X and Y that the increment check, at least for a good forecast model, is nearly as accurate as the horizontal check. With these facts in mind, it was decided to make the code so that it would operate well without the availability of the horizontal check. This would allow efficient testing, on workstation or even PC, of the DMA.

Furthermore, it is desirable to calculate the horizontal check only once, and not on each scan. In order to avoid recalculating the horizontal residual, it is necessary to be sure that the value of the data at influencing stations would not change as a result of qc actions, i.e. that they are good from the beginning. Therefore, the following procedure is made: 1) calculate all residuals, except the horizontal residual, 2) make a preliminary quality judgment based on these residuals, 3) calculate the horizontal residuals, excluding from use any data that receive a preliminary judgment of bad. With this procedure, any correction to a value during the qc scans will have a simple effect on the horizontal residual so that it need not be recalculated from scratch.

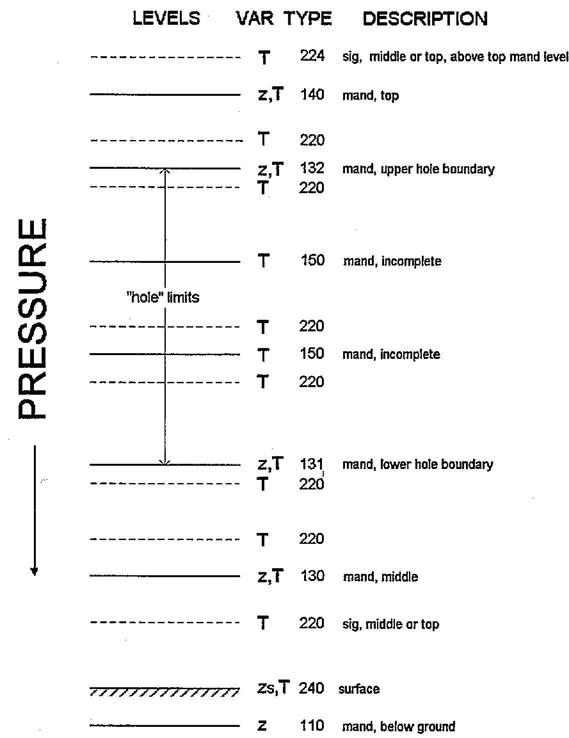

Fig. 14 Assignment of level types, below ground, surface, and mandatory and significant levels above ground.

The logic used by the DMA is determined by which checks are available and relevant, and this is, in turn, determined by the type of level. Fig. 14 shows the various level types. A statement about each type is made:

110 - mandatory, below ground
Sometimes there are mandatory levels reported below the ground level. They are obtained by hydrostatic integration, generally at the reporting station, to these levels with an assumed temperature profile. Sometimes, these levels may be used indirectly by NCEP operations, so they are qc'd by looking for observation errors in them.

240 - surface
The surface is a special significant level in that both height and temperature are reported at this level, making it in effect a mandatory level. It is also one boundary for baseline computations.

120 - mandatory, first above ground
This level is special because it uses the baseline residual below it, but the regular hydrostatic residual above.

220 - significant, middle or top
This is the usual significant level. The vertical and increment checks are most powerful. No correction will be made unless there is evidence of a large hydrostatic residual in the layer containing this level.

131 - mandatory, middle, lower hole boundary
A hole refers to a significant pressure interval through which data are not available for accurate hydrostatic calculations. It is specifically defined as a layer, containing at least two intermediate mandatory levels which do not have complete data, i.e. height or temperature missing. Surface type 131 refers to the lower (in height) boundary of a hole.

150 - mandatory, incomplete
This is a mandatory level with height or temperature missing. Two or more adjacent surface type 150 levels are contained within a hole.

132 - mandatory, middle, upper hole boundary
The upper boundary of the hole, corresponding to surface type 131.

130 - mandatory, middle
The usual mandatory level, containing non-missing height and temperature.

133 - mandatory, middle, isolated
It is possible that a mandatory level have a hole both above and below. In such a case, the surface type is 133. For such a level, the useful residuals are limited to increment and horizontal.

140 - mandatory, top
The top mandatory level is special because a hydrostatic residual is only available below.

224 - significant, middle or top, above top mandatory level
At such a level, no layer hydrostatic residual is available. Therefore, only observation errors can be examined; no corrections are warranted.

The first step in the DMA is to determine the surface type. There is specific logic for each surface type. This logic will be discussed in following subsections. However, there are a limited number of specific errors that are considered and even fewer routines that determine them. Table 1 lists the specific error types, giving an error type number for each.

Table 1. Errors Detectable by Complex Quality Control for Rawinsonde Heights and Temperatures

Type	Description
	Communication and Computation Errors
	Errors at a Single Interior Mandatory Level
1	Single height
2	Single temperature
3	Height and temperature at the same level
	Height and temperature at the same level with residual compensation
	Errors at the Top Mandatory Level
5	Height, temperature, or both
	Computation Error in Height at Mandatory Levels
6	Height computation between any two mandatory levels
	Errors at Adjacent Mandatory Levels
7	Height at two adjacent levels
8	Temperature at two adjacent levels
9	Height at the lower, and temperature at the upper of two adjacent levels
10	Temperature at the lower, and height at the upper of two adjacent levels
	Errors at Significant Levels
20	Significant level temperature corrected
21-25	Non-correctable significant level temperature errors
	Surface Errors
100	Surface pressure communication error
102	Surface temperature error
105	Likely surface temperature error, too small to correct
	Undetermined error(s), possibly in surface pressure
106	Surface pressure observation error
	Observation Errors
30-35	Temperature observation errors, rejected or used with reduced weight
36,37	Height observation errors, rejected or used with reduced weight

Each group of errors in the table, separated by a horizontal line, has a separate routine for its diagnosis. It is convenient to introduce their names now as a shorthand notation for future reference. For errors at a single interior mandatory level, types 1, 2, and 3, the routine called is **ERR123**. For errors at the top mandatory level, type 5, **ERR5** is called. For computation errors, type 6, **COMPER** is called. For errors at adjacent mandatory levels, types 7-10, the routine **ERR710** is called. For errors at a significant level, types 20-25, the routine **SIGERR** is called. And for observation errors, types 30-37, **OBERR** is called. The details of each of these routines will be outlined after consideration of the logic associated with each surface type.

Each surface type may potentially have errors in several groups. From the magnitudes of the hydrostatic residuals available, which errors to consider is determined. The following sub-sections discuss this in more detail.

a. Surface type 130 - mandatory, middle

The examination begins with the hydrostatic residuals. For a surface type 130, the hydrostatic residuals are available both above and below, and may also be available at the next level below as well. Each of these 3 residuals is examined to see if it is large, with subsequent action depending upon the result. The code decides between possible observation errors, errors at a single level, errors at multiple levels, or a computation error. Each of these error types is represented by a subroutine where the actual checking and possible correction is performed. However, since computation errors are checked on a separate, last scan through the data, they are not checked here.

b. Surface type 120 - first mandatory level above the surface

The only way that the first mandatory level differs from any middle mandatory level is that the hydrostatic residual for the layer below the data level is replaced by the baseline residual. Otherwise the logic remains the same. Errors may be observation (considered on the last scan, separately), errors to one or more values at 1 or 2 levels, or computation errors.

c. Surface type 240 - earth's surface

The earth's surface is special because of the use of the baseline check and because of the multiplicity of kinds of errors that are possible. There can be a surface pressure communication error, surface pressure observation error, surface temperature error, or undetermined errors. In addition, there can be an error in the station elevation, but this can only be established by a history of the baseline residual for a month or more.

Each error at the surface has characteristics that distinguish it from others. The objective is to choose the type that best fits the residuals. Table X shows some basic information on the various residuals for the different surface errors. The error is assumed to have the magnitude p.

Type of error	BASRES		PSINC		PIS
p_s - communication	p·(dz/dp)		p > PSLIM		p
p_s - observation	0.		0.		p > PSLIM
Ts	z > 20.		z/(dz/dp)		0.
Other	> 20.	or	> PSLIM	or	> 2·PSLIM

In the table, dz/dp is the vertical derivative of height with respect to pressure. As seen below, it is estimated hydrostatically. PSLIM is a constant with value 3.0 hPa.

In the table, BASRES is the baseline residual, PSINC is the baseline residual in terms of pressure, and PIS is the surface pressure increment. A further description of the conditions that distinguish between the surface errors is given below.

1. Surface pressure communication error (type 100)

For a surface pressure communication error, the baseline residual in terms of pressure (PSINC) and the surface pressure increment (PIS) should be close to each other. Also, the closeness should be relaxed somewhat for larger PSINC. These factors contribute to the definition of R1, which should be small for a surface pressure communication error. If |PSINC| > PSLIM and PSINC ≠ missing, then calculate R1:

$$R1 = \frac{|PSINC - PIS|}{PSLIM \cdot C1 \cdot \ln(|PSINC|)}$$

If R1 < 1., then the error could possibly be a surface pressure communication error, but other possibilities must be checked as well.

2. Surface pressure observation error (type 106)

For the diagnosis of a surface observation error, there are two factors that must be taken into account. First, there must be agreement between the surface pressure increment and the height increments. And second, the height increments must be consistent vertically. Naturally, the increments must be large enough in the first place to indicate that there is an error.

For a surface observation error, the baseline residual (BASRES) is small, while the surface pressure increment (PIS) is large. As indicated above, PIS should correspond to the increment of the first mandatory level height, with opposite sign. To make the values comparable, the height must be divided by dz/dp, the derivative of height with respect to pressure. It is approximated by

$$\frac{dz}{dp} = \frac{R}{g}\frac{T_c + T_0}{p_{LM}}$$

where the variables have the same meanings as before and p_{LM} is the pressure at the first mandatory level above the ground.

If |PIS| > PSLIM (indicating an error) and PIS is non-missing, ~~and |BASRES| > 2.0 and BASRES is non-missing,~~ then calculate R2 and R3:

$$R2 = \frac{\left| PIS - ZI_{LM} \Big/ \frac{dz}{dp} \right|}{PSLIM}$$

$$R3 = \frac{SX \Big/ \frac{dz}{dp}}{PSLIM}$$

where ZI_{LM} is the height increment at the lowest complete mandatory level and SX is the standard deviation of the height increment at all levels. R2 measures the agreement between the surface pressure increment and the height increments, while R3 measures the consistency of the height increments themselves. An additional measure of the height increment agreement, R5, is also calculated:

$$R5 = \frac{\left| ZI_{LM+1} - ZI_{LM} \right| \Big/ \frac{dz}{dp}}{PSLIM}.$$

Whichever of R2, R3/0.65, and R5 is largest, if it is less than 1.0, is used to compare against the other possible surface error types.

3. Surface temperature communication error (type 102)

When there is a surface temperature communication error, the surface pressure increment is small, but the baseline residual is large. A preliminary temperature correction, TSCOR, is calculated based upon BASRES. This estimated correction should be close to a weighted average of the other temperature residuals, TCBEST. This is expressed by R4, defined below, being small. If |BASRES| > 20 and BASRES is non-missing, then calculate TSCOR and R4 as

$$TSCOR = -\frac{BASRES}{\frac{R}{2g}\ln\left(\frac{p_{LM+1}}{p_{LM}}\right)}$$

$$R4 = \frac{|TSCOR - TCBEST|}{XINC}$$

If R4 < 0.25, then a surface temperature communication error is possible.

The final selection of the most likely surface error is made on the basis of whichever of R1, maximum of (R2,R3/0.65,R5), or R4/0.25 is the smallest and less than 1.0. Otherwise, if there is still found to be an error, an undefined type (105) is assigned.

d. Surface type 140 - top mandatory level, and surface type 131 - lower hole boundary

The logic for these types is identical; neither has a hydrostatic residual above that can be used for error determination. When the hydrostatic residual for the layer below the data level is large, then **ERR5** is called. As usual, the residual is called large when it exceeds 3.5 times the standard deviation of the residual, determined for an extended period of time when there is no error.

e. Surface type 132 - middle mandatory level, upper hole boundary, and top mandatory level, upper hole boundary, and isolated middle mandatory level

Since the hydrostatic residual for the layer below would extend over a very large pressure interval, it is unreliable. Therefore, no correction of data is attempted for data for these surface types.

f. Surface type 121 - first mandatory level above ground, lower hole boundary

The only relevant hydrostatic residual is the baseline residual. If it is large then errors at the single data level are considered by calling **ERR123**.

g. Surface type 220 - middle significant level

At a significant level, the hydrostatic residual for the layer containing the level cannot often be used to determine the magnitude of an error. It is used in the weaker way by cqcht96 to indicate whether or not there is a communication error. The effect of a significant level temperature error upon the hydrostatic residual will extend only from the level upward and downward to the next significant or mandatory level. For significant levels that are close together, this effect may be quite small, even for a rather large temperature error. In order to consider the possible effect of an error, the hydrostatic residual is scaled by the logarithm of the pressure thickness. If the scaled value is large enough, then a communication error is further considered in **SIGERR**.

h. Surface type 244 - significant level, above top mandatory level

For a significant level above the top mandatory level, there is no hydrostatic residual available, and thus it is impossible to know whether a communication error is present. On the last scan, a search for observation errors will be made.

i. Surface type 150 - incomplete mandatory level

If the height is missing from a mandatory level, then the information available is the same as at a significant level, except that a horizontal residual may be also available. At present, a surface type 150, which is a rare type, is treated the same as a significant level.

7. Detail of error correction routines

There is a unique routine for each of the error classes as described before. Each of these routines may be called as appropriate for more than one surface type. There are some general principles followed in these routines that will be outlined first, thereby avoiding repetition.

The first attempt at error correction uses the hydrostatic residual(s) to make a suggested value. For a height correction, the most accurate hydrostatic residual is calculated using significant as well as mandatory level heights, while for a temperature correction, the most appropriate hydrostatic residual is calculated with only mandatory level data, allowing the error to be "felt" over the whole layer.

If an error is caused by human error, as most are, then the hydrostatically proposed error should be close to the actual error. Therefore, it is appropriate to look within the vicinity of this proposed correction for a "simple" correction. Such a correction is one that would lead to a change in sign, single digit, interchange of digits, or a combination of these changes. In many cases a simple correction is found, and it forms the basis for further testing. It is first rounded as appropriate for the data type and level. The pressure is rounded to the nearest 0.1 hPa, height is rounded to the nearest meter up to 700 hPa and to the nearest 10. meters above, and the temperature is rounded to the nearest 0.1 degree.

A second proposed correction is formed as a linear combination of the available residuals, weighted inversely as the long-term standard deviation of each residual in the absence of error. If the hydrostatically proposed correction is close enough to this "best" correction, then it is accepted. If it is not close enough, then the "best" correction itself forms the basis for the next trial solution. Near this value a "simple" correction is sought, and the value is rounded. This proposed correction will be accepted if it makes all the available residuals smaller. If not, then the original "best" correction will likewise be tested. In none of these corrections is good, then the value will be flagged as bad.

a. Corrections to single height, single temperature, or height and temperature at the same (mandatory) level

When hydrostatic residuals are large at two consecutive levels, the corrections to single height, single temperature, or height and temperature at the same (mandatory) level are appropriate. Cqcht96 calls the subroutine ERR123 to make the diagnosis. Three quantities are calculated that help to decide between the possible error types. They are:

$$H_1 = S^s_{L1} + S^s_{L2}$$

$$H_2 = 2\Delta\sqrt{B^2_{L1} + B^2_{L2}}$$

$$T_1 = \sqrt{\frac{S^m_{L2}}{S^m_{L1}}}$$

where *L1* stands for the layer below the data level and *L2* stands for the layer above the data level. The level L1 may in fact extend to two levels below if the necessary data are missing at the first level below. The constant Δ has the value 7.0. Remember that S^s is the hydrostatic residual calculated using all levels, while S^m is the hydrostatic residual calculated only mandatory levels.

A height error is diagnosed if $H_1 < H_2$ and $T_1 > \Delta$ and the residuals S^m_{L1} and S^m_{L2} are both greater than 3.5 times their historical standard deviation. The hydrostatically suggested height correction is:

$$ZCOR = \frac{\frac{S^s_{L2}}{B^2_{L2}} - \frac{S^s_{L1}}{B^2_{L1}}}{\frac{1}{B^2_{L2}} + \frac{1}{B^2_{L1}}}.$$

From this point, the analysis proceeds as discussed in the introductory part of this section.

A temperature error is diagnosed if $T_1 < \Delta$ and $H_1 \geq H_2$. The hydrostatically suggested temperature correction is:

$$TCOR = \frac{1}{2}\left(\frac{S^m_{L1}}{B_{L1}} + \frac{S^m_{L2}}{B_{L2}}\right).$$

And the analysis proceeds from this value.

If the hydrostatic residuals S^m_{L1} and S^m_{L2} are both less that 3.5 times their historical standard deviation, then there is no error at all. Otherwise, errors to both height and temperature are sought, with the hydrostatically determined corrections defined by:

$$ZCOR = \frac{B_{L2}S^s_{L2} - B_{L1}S^s_{L1}}{B_{L1} + B_{l2}}$$

$$TCOR = \frac{S^m_{L1} + S^m_{L2}}{B_{L1} + B_{L2}}$$

The principles outlined in the introduction to this section are followed in determining just what, if any, correction is to be made. The number of possibilities is larger because of the attempt to correct two values simultaneously.

b. Corrections at the top mandatory level

At the top mandatory level, or at the lower boundary of a hole, only the hydrostatic residual for the layer below may be used in assisting to correct a height, temperature, or both. The subroutine, ERR5, first tries a height correction. If that fails, it next tries a temperature correction. And if that fails, it tries a correction to both, where the trial corrections are the "best" corrections. If no correction is possible, then the data are appropriately flagged.

The hydrostatically suggested height and temperature corrections are given by:

$$ZCOR = -S_{L1}^{s}$$

$$TCOR = \frac{S_{L1}^{m}}{B_{L1}}.$$

c. Corrections to adjacent levels: 2 heights, 2 temperatures or 1 height and 1 temperature

When there are either three large hydrostatic residuals at consecutive levels or two large hydrostatic residuals, separated by a smaller one, then there may be errors in height and/or temperature at two adjacent mandatory data levels. Four specific error types are examined: type 7—error in height at the two levels, type 8—error in temperature at the two levels, type 9—error in the lower height and upper temperature, and type 10—error in the lower temperature and upper height.

The diagnosis for these error types is rather complicated and not particularly fruitful. It stretches the limits of what can reasonably be done with cqcht, given the nature of the data and its random and sampling error sources.

The diagnosis begins by defining four quantities, each of which must be small for the corresponding error type to be likely. A derivation of these quantities will not be given. The requirement that these quantities be small can be considered to be existence conditions. The quantities are:

$$R7 = \frac{\left|S_{L1} + S_{L2} + S_{L3}\right|}{\Delta\sqrt{B_{L1}^2 + B_{L2}^2 + B_{L3}^2}}$$

$$R8 = \frac{\left|\frac{S_{L1}}{B_{L1}} - \frac{S_{L2}}{B_{L2}} + \frac{S_{L3}}{B_{L3}}\right|}{\sqrt{3}\Delta}$$

$$R9 = \frac{\left|S_{L1} + S_{L2} - \left(\frac{B_{L2}}{B_{L3}}\right)S_{L3}\right|}{\Delta\sqrt{B_{L1}^2 + 2B_{L2}^2}}$$

$$R10 = \frac{\left|S_{L1} + S_{L3} - \left(\frac{B_{L2}}{B_{L1}}\right)S_{L1}\right|}{\Delta\sqrt{B_{L3}^2 + 2B_{L2}^2}}$$

where all hydrostatic residuals use mandatory level information only, i.e. all S are S^m. Other quantities have the same meaning as before.

Combinations of the residuals must be of sufficient magnitude for errors to be suspected. The following logical quantities are used in assessing the necessary magnitude:

$$C2 = \text{ true if } |S_{L1}| > \Delta\sqrt{B_{L1}^2 + 2B_{L2}^2}$$

$$C3 = \text{ true if } |S_{L3}| > \Delta\sqrt{B_{L3}^2 + 2B_{L2}^2}$$

$$C5 = \text{ true if } \left|\frac{S_{L1}}{B_{L1}}\right| > \Delta$$

$$C6 = \text{ true if } \left|\frac{S_{L3}}{B_{L3}}\right| > \Delta$$

The most likely error type is determined from the following conditions.

Type 7 if R7 < 1 and C2 is true and C3 is true.
Type 8 if R8 < 1 and C5 is true and C6 is true.
Type 9 if R9 < 1 and C2 is true and C6 is true.
Type 10 if R10 < 1 and C3 is true and C5 is true.

The suggested corrections are given by the following formulas. For type7—errors of two consecutive heights—the suggested corrections are:

$$ZCOR_{I1} = \frac{-\dfrac{S_{L1}}{B_{L1}^2}\left(\dfrac{1}{B_{L2}^2} + \dfrac{1}{B_{L3}^2}\right) + (S_{L2} + S_{L3})\dfrac{1}{B_{L2}^2 B_{L3}^2}}{\left(\dfrac{1}{B_{L1}^2} + \dfrac{1}{B_{L3}^2}\right)\dfrac{1}{B_{L2}^2} + \dfrac{1}{B_{L1}^2 B_{L3}^2}}$$

$$ZCOR_{I2} = \frac{-(S_{L1} + S_{L2})\dfrac{1}{B_{L1}^2 B_{L2}^2} + \dfrac{S_{L3}}{B_{L3}^2}\left(\dfrac{1}{B_{L1}^2} + \dfrac{1}{B_{L2}^2}\right)}{\left(\dfrac{1}{B_{L1}^2} + \dfrac{1}{B_{L3}^2}\right)\dfrac{1}{B_{L2}^2} + \dfrac{1}{B_{L1}^2 B_{L3}^2}}$$

where the levels I1 and I2 are the interfaces between the layers L1, L2 and L3.

For type 8—errors of two consecutive temperatures—the suggested corrections are:

$$TCOR_{I1} = \frac{1}{3}\left(2\frac{S_{L1}}{B_{L1}} + \frac{S_{L2}}{B_{L2}} - \frac{S_{L3}}{B_{L3}}\right)$$

$$TCOR_{I2} = \frac{1}{3}\left(2\frac{S_{L3}}{B_{L3}} + \frac{S_{L2}}{B_{L2}} - \frac{S_{L1}}{B_{L1}}\right).$$

For type 9—errors of the lower height and upper temperature—the suggested corrections are:

$$ZCOR_{I1} = \frac{\frac{S_{L2}}{B_{L2}^2} - 2\frac{S_{L1}}{B_{L1}^2} - \frac{S_{L3}}{B_{L2}B_{L3}}}{\frac{2}{B_{L1}^2} + \frac{1}{B_{L2}^2}}$$

$$TCOR_{I2} = \frac{\frac{S_{L2}}{B_{L2}B_{L1}^2} + \frac{S_{L1}}{B_{L1}^2 B_{L2}} + \frac{S_{L3}}{B_{L3}}\left(\frac{1}{B_{L1}^2} + \frac{1}{B_{L2}^2}\right)}{\frac{2}{B_{L1}^2} + \frac{1}{B_{L2}^2}}.$$

And for type 10—errors of the lower temperature and upper height—the suggested corrections are:

$$TCOR_{I1} = \frac{\frac{S_{L1}}{B_{L1}}\left(\frac{1}{B_{L2}^2} + \frac{1}{B_{L3}^2}\right) + \frac{S_{L3}}{B_{L2}B_{L3}^2}}{\frac{1}{B_{L2}^2} + \frac{2}{B_{L3}^2}}$$

$$ZCOR_{I2} = \frac{\frac{S_{L1}}{B_{L1}B_{L2}} + 2\frac{S_{L3}}{B_{L3}^2} - \frac{S_{L2}}{B_{L2}^2}}{\frac{1}{B_{L2}^2} + \frac{2}{B_{L3}^2}}.$$

The suggested corrections are tested for acceptability and the corrections are made or the data are flagged.

d. Computation error corrections and errors to height and temperature at the same level with hydrostatic residual compensation

When there is a single large hydrostatic residual, there are two possible causes. The most common reason is a computation error in working up the heights. In this case, all heights above the computation error are in error by the same amount, namely the negative of the hydrostatic residual. Therefore, for a computation error, the correction for all heights above is:

$$ZCOR = -S^{s}_{L1}.$$

In the special case where the computation error takes place between the surface and the first mandatory level, then the hydrostatic residual is replaced by the baseline residual.

Unfortunately, there is a second possible reason for an isolated large residual. When there is an error to both height and temperature at the same level, then there is the possibility for the hydrostatic residuals to become small at one of the two levels. Allowing for that possibility greatly complicates the situation. Since that is a rather rare occurrence, the details will not be given. In brief, the code decides which type of error is most likely and then applies the tentative correction. For a computation error, the correction is not accepted unless the hydrostatic residual is made smaller for several levels above. For the "type 3 with compensation" error, each or both corrections may be accepted or the data flagged.

e. Errors in temperature at significant levels

No correction is attempted at a significant temperature level unless there is evidence from the hydrostatic residual, calculated for the layer containing the data level, that there was a communication error. This is so because it would be inappropriate to make changes to temperatures containing observation errors. Beyond this, however, the hydrostatic residuals are not used for significant level diagnosis. The residuals that can be used are the increment and the vertical residual.

The various residuals are scaled by their historical standard deviations to give numbers NTI for the increment, NTV for the vertical residual, NTVP for the vertical residual at the next level above, and NTVM for the vertical residual at the next level below. The analysis of significant level errors depends upon the value of these quantities. There are 5 possibilities considered as described in the rest of this section.

When NTI and NTV are both > 3.5, then an error is definitely present and a correction is attempted. If the increment, TI, and vertical residual, TV, are close to each other, then the "best" correction will be attempted. As usual, a simple correction is sought, and it is rounded before checking. If this proposed correction is acceptable, it will be used. If TI and TV are not close enough to each other, then -TI will be tried in a similar way as a correction. In either case, if the correction is not accepted, then the data quality is set to bad.

When NTI > 3.5 (but NTV ≤ 7) then the increment is large, but, since there is some measure of vertical consistency, it would be dangerous to make a correction. The data are marked as questionable.

It is desired to avoid flagging or "correcting" the Tropopause level temperature. At such a level, the vertical residual may be large, but the vertical residual should be small at neighboring levels. Therefore, when NTV > 7 and NTVP < 5 and NTVM < 5, the quality is set to good.

In a little less clear case, when NTV > 7 and NTVP < 7 and NTVM < 7, the temperature at the data level stands out, but not as much as at the Tropopause and not enough for the increment to be large. In this case, the data quality is set to questionable.

In all other cases, the data quality is set to good. The reason for the apparent influence of bad data upon the layer hydrostatic residual must be found at other levels.

f. Observation errors

The term "observation error" is used to signify errors from several sources. These errors are looked for only after all correctable errors have been considered and mandatory and significant levels are examined together.

One possible source is unresolved cases of large hydrostatic errors, where moderate height or temperature errors are not corrected or flagged. Such errors should lead to other residuals of moderate magnitude. They will generally be isolated.

There may be an isolated observation error in temperature with undetermined cause but moderate to large magnitude. Some possible reasons would be temporary icing on the instrument, a communications glitch between the sonde and the ground, etc. In such cases the hydrostatic residuals are small, the lapse rate may be unstable, and other residuals, including the vertical residual, are large.

In many other cases, there are several levels with small to moderate temperature errors. In such cases it may be difficult to tell over what layers the errors extend. In such cases, the hydrostatic residuals are small, the increments and/or horizontal residuals are large, the vertical residuals are generally small, and the lapse rates are stable. These small to moderate temperature errors accumulate hydrostatically to cause height observation errors. They are detected most clearly by the increment and horizontal checks.

With these several causes for observation errors, the detection is complicated. The diagnosis begins by calculating the "best" residuals for height and temperature: linear combinations of the available residuals. The requirements for error detection begin by testing these "best" values against various criteria.

Temperature observation errors are considered first. The diurnal variation and the large gradients of temperature near the earth's surface make error determination particularly difficult here. For that reason, the required magnitude of the "best" residuals for errors to be diagnosed is expanded near the surface. The factor by which they are expanded is shown in Fig. 15. In the following, this value is referred to as CON.

There are a number of factors that are used in determining the presence of a temperature observation error. They are: 1) the magnitude of the "best" residual at the data location, whether small, moderately large, or large, 2) the magnitudes of the "best" residual and vertical residuals at adjacent levels (telling whether the error is isolated), and 3) the lapse rates above and below the data level. For temperature, the "best" residual is moderately large if it exceeds 0.293 * CON * temperature increment historical standard deviation, and large if it exceeds 0.440 * CON * temperature increment historical standard deviation.

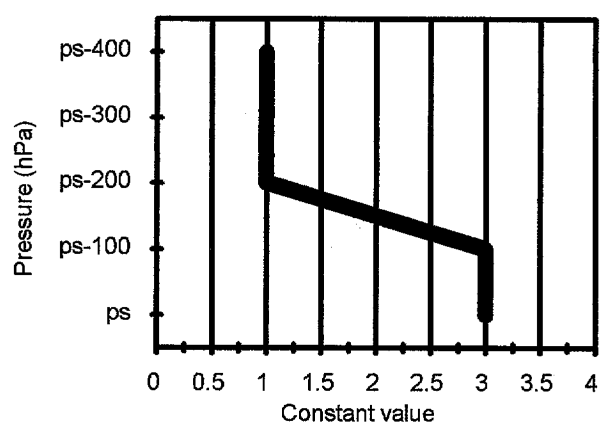

Fig. 15 The variation of the constant CON with difference in pressure from the surface pressure.

The determination of height observation error is simpler. First, the "best" residual is calculated. If it exceeds 0.605 * height increment historical standard deviation, then the height is flagged as bad, and if it exceeds 0.403 * height increment historical standard deviation, then the height is flagged as questionable.

8. Output produced by CQCHT96

The primary output from cqcht96 is a BUFR file of special format, called prepbufr at NCEP. This file contains not only the initial report but a record of all changes to the data and quality assessments. This file is used later by the Optimal Interpolation QC data quality control program, which quality controls several data types without the possibility of correction) and as input to the analyses.

For monitoring purposes, there is a print file that is produced by running cqcht96. For any reports with errors that are diagnosed, this file contains the report, all residuals, a diagnosis of any errors and corrections, the report after corrections if there are any, and a summary of the information that will be placed in the prepbufr file—the so-called events. There is also a statistical summary of all residuals for all reports received and summary lists of the events. The information for the examples in the next section were taken from this print file. However, the examples will show only the necessary information. The complete information is displayed in Fig. 16.

The first part of the printout is called the quick recognition part. It displays information on the residuals and diagnoses in abridged form. The residuals are grouped by height, temperature and dew-point temperature. The numbers below are the absolute value of the residuals, scaled so that 20 corresponds to 10 standard deviations for the historical record of the residual. The various residuals are I - increment, V - vertical, H - horizontal, HY - hydrostatic, and LP - lapse rate. The value -1 signifies missing. Toward the right are error types for each variable, P - pressure, Z - geopotential height, T - temperature, and Td - dew-point temperature, and for each level. The normalized baseline (BAS) and pressure increment (IPSINC) residuals are at the far right. The pattern of residuals may be clearly seen from this table to quickly assess errors. More complete analysis requires looking at the next part.

Below the quick recognition part is the body of the diagnosis. It shows complete information on the report, including all residual values, the surface types and category for each level. At the top is information about the station location, date/time information, instrument type, surface pressure, and baseline residuals. Below, by pressure, are the values of height, temperature, dew-point temperature, and the residuals for increment, vertical, horizontal, and hydrostatic. The hydrostatic residual, scaled to temperature, X, is also given. In this part, missing values are signified by asterisks. The quick recognition part and this part are repeated after the list of events when there are corrections. In the example shown here, there were corrections, but the parts are not repeated. Note that the lapse rate classes are given only in the quick recognition part.

The third part shows a list of the events that will be included in the prepbufr file. It is possible that there is more than 1 diagnosis or correction for the same datum. This list compresses the events so that there is at most 1 event for each datum. A complete list of these events for all stations is given at the end of the printout. There is also a list there of all events, including possibly more than 1 for the same datum. The list of events in Fig. 16 shows the pressure, variable (P - pressure, T - temperature, or Z - height), quality mark (1 - good, 3 - questionable, or 13 - bad), original value, correction, new value, level count (including all levels), station identifier, station sequence number in report (SQN), and date/time of the report.

There are three additional files that are produced that make summaries at periodic intervals, usually monthly, possible. One file contains the count of the number of stations received for each WMO station block. A second file contains information on each events, similar to that available within the print file. And a third file contains information about each report received: number of levels of each category (mandatory, significant, etc.) and the increments of height, temperature and specific humidity at each level. Many interesting uses may be made of this information. Section 10 shows overall performance of cqcht96 as extracted from these results.

STN ID: 89664 LAT: -77.85 LON: 166.67 STN HT: 24.

	--Height--				----Temp-----					-DewP T-			------IHSC------					
PRESS	I	V	H	HY	I	V	H	LP	HY	I	V	H	P	Z	T	Td	BAS	IPSINC
1000.0	16	12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	36	0	0	20	4
974.0	15	17	-1	-1	5	6	-1	1	-1	2	10	-1	100	0	0	0		
973.0	-1	-1	-1	-1	11	19	-1	4	-1	2	3	-1	0	36	0	0		
968.0	-1	-1	-1	-1	5	3	-1	3	-1	5	16	-1	0	36	0	0		
957.0	-1	-1	-1	-1	5	2	-1	1	-1	4	9	-1	0	0	31	0		
934.0	-1	-1	-1	-1	6	7	-1	4	-1	6	14	-1	0	0	31	0		
925.0	0	20	-1	4	1	5	-1	1	0	2	15	-1	0	0	0	0		
864.0	-1	-1	-1	-1	2	7	-1	4	-1	4	14	-1	0	0	0	0		
850.0	0	1	-1	20	3	11	-1	1	2	7	18	-1	0	0	0	0		
840.0	-1	-1	-1	-1	4	9	-1	1	-1	5	12	-1	0	0	0	0		
747.0	-1	-1	-1	-1	7	6	-1	1	-1	5	2	-1	0	0	0	0		
706.0	-1	-1	-1	-1	6	12	-1	4	-1	3	11	-1	0	0	0	0		
700.0	4	3	-1	15	4	19	-1	1	0	9	20	-1	0	0	0	0		
638.0	-1	-1	-1	-1	8	20	-1	3	-1	8	20	-1	0	0	0	0		
500.0	6	1	-1	3	3	7	-1	1	3	9	20	-1	0	0	0	0		
400.0	8	7	-1	0	6	6	-1	1	4	-1	-1	-1	0	0	0	0		
349.0	-1	-1	-1	-1	5	5	-1	1	-1	-1	-1	-1	0	0	33	0		
348.0	-1	-1	-1	-1	2	1	-1	1	-1	-1	-1	-1	0	0	0	0		
312.0	-1	-1	-1	-1	0	1	-1	1	-1	-1	-1	-1	0	0	0	0		
300.0	9	10	-1	0	2	2	-1	1	1	-1	-1	-1	0	0	0	0		
291.0	-1	-1	-1	-1	3	0	-1	1	-1	-1	-1	-1	0	0	0	0		
250.0	6	6	-1	2	6	8	-1	1	0	-1	-1	-1	0	0	0	0		
201.0	-1	-1	-1	-1	3	0	-1	1	-1	-1	-1	-1	0	0	0	0		
200.0	2	0	-1	1	5	4	-1	1	1	-1	-1	-1	0	0	33	0		
150.0	0	2	-1	0	1	1	-1	1	0	-1	-1	-1	0	0	0	0		
100.0	0	0	-1	1	1	3	-1	1	1	-1	-1	-1	0	0	0	0		
70.0	1	1	-1	0	0	0	-1	1	0	-1	-1	-1	0	0	0	0		
50.0	1	2	-1	-1	1	1	-1	-1	-1	-1	-1	-1	0	0	0	0		

STN ID: 89664 LAT: -77.85 LON: 166.67 STN HT: 24.
DATE/TIME: 97041600 DHOUR: 0.0 SCAN: 1 INST TYPE: 43
SURFACE PRESS: 974.0 PIS: -19.3 PSINC: -17.3 BASRES: 133.

	<--------Height---------->				<-Hydr Res-->		<-----------Temperature---------->					<--Dew Point Temperature-->					
PRESS	ZOB	ZI	ZV	ZH	HYDS	HYDN	TOB	TI	TV	TH	X	TDO	TDI	TDV	TDH	LST	CAT
1000.0	-179.	-153.	-58.	******	******	******	******	******	******	******	******	******	******	******	******	110	1.
974.0	24.	-145.	-82.	******	******	******	-14.1	7.3	-3.1	******	******	-22.1	3.0	5.4	******	240	0.
973.0	******	******	******	******	******	******	-5.8	15.7	9.7	******	******	-28.7	-3.6	-1.7	******	220	2.
968.0	******	******	******	******	******	******	-14.1	7.7	-1.6	******	******	-33.1	-7.7	-8.3	******	220	2.
957.0	******	******	******	******	******	******	-15.1	7.2	1.2	******	******	-21.1	5.2	5.1	******	220	2.
934.0	******	******	******	******	******	******	-15.9	7.4	4.1	******	******	-18.7	8.3	7.8	******	220	2.
925.0	547.	0.	62.	******	-8.	0.	-21.7	1.7	-2.8	******	-0.1	-30.7	-3.6	-8.6	******	120	1.
864.0	******	******	******	******	******	******	-19.9	2.6	3.9	******	******	-20.4	3.9	8.1	******	220	2.
850.0	1163.	-4.	5.	******	-50.	-3.	-26.7	-3.5	-6.2	******	-1.2	-31.7	-6.8	-10.4	******	130	1.
840.0	******	******	******	******	******	******	-19.5	4.2	5.1	******	******	-20.3	5.1	6.8	******	220	2.
747.0	******	******	******	******	******	******	-22.9	5.9	2.7	******	******	-25.6	4.1	0.9	******	220	2.
706.0	******	******	******	******	******	******	-26.3	4.8	5.0	******	******	-28.2	3.1	4.7	******	220	2.
700.0	2537.	-26.	-8.	******	-49.	0.	-34.9	-3.5	-7.7	******	-0.1	-39.0	-7.4	-11.1	******	130	1.
638.0	******	******	******	******	******	******	-28.1	7.2	9.1	******	******	-28.7	6.6	12.4	******	220	2.
500.0	4830.	-51.	-6.	******	-8.	-8.	-45.7	-2.6	-2.9	******	-2.3	-50.7	-7.5	-11.2	******	130	1.
400.0	6280.	-86.	-27.	******	-1.	-15.	-54.3	-5.4	-3.1	******	-3.7	******	******	******	******	130	1.
349.0	******	******	******	******	******	******	-58.1	-5.0	-3.0	******	******	******	******	******	***********		5.
348.0	******	******	******	******	******	******	-55.1	-2.0	1.0	******	******	******	******	******	******	220	2.
312.0	******	******	******	******	******	******	-55.7	-0.3	-0.9	******	******	******	******	******	******	220	2.
300.0	8110.	-109.	-44.	******	-3.	-3.	-53.7	2.3	1.2	******	-1.1	******	******	******	******	130	1.
291.0	******	******	******	******	******	******	-53.1	2.9	0.4	******	******	******	******	******	******	220	2.
250.0	9290.	-85.	-27.	******	11.	3.	-49.3	6.7	4.9	******	1.0	******	******	******	******	130	1.
201.0	******	******	******	******	******	******	-50.3	3.9	-0.7	******	******	******	******	******	******	220	2.
200.0	10760.	-37.	-4.	******	5.	5.	-47.9	6.2	3.5	******	1.2	******	******	******	******	130	1.
150.0	12660.	4.	20.	******	-2.	-2.	-48.3	2.3	1.3	******	-0.3	******	******	******	******	130	1.
100.0	15310.	-2.	8.	******	-9.	-9.	-51.1	-2.2	-2.4	******	-1.7	******	******	******	******	130	1.
70.0	17610.	-31.	-14.	******	-1.	-1.	-52.9	-0.9	-0.2	******	-0.3	******	******	******	******	130	1.
50.0	19760.	-43.	-27.	******	******	******	-56.5	-1.4	-1.2	******	******	******	******	******	******	140	1.

PRESSURE	VAR	IETYP	QMARK	ORIG-VAL	COR	NEW-VAL	LEVEL	LEVTYP	STATION	SQN	DATE
974.0	P	100	1.	974.0	17.3	991.3	2	1	89664	2.	97041600
638.0	T	20	1.	-28.1	-10.0	-38.1	14	4	89664	2.	97041600
1000.0	Z	36	13.	-179.0	0.0	-179.0	1	3	89664	2.	97041600
747.0	T	31	3.	-22.9	0.0	-22.9	11	4	89664	2.	97041600
706.0	T	31	3.	-26.3	0.0	-26.3	12	4	89664	2.	97041600
400.0	T	32	3.	-54.3	0.0	-54.3	16	2	89664	2.	97041600
400.0	Z	37	3.	6280.0	0.0	6280.0	16	2	89664	2.	97041600
300.0	Z	37	3.	8110.0	0.0	8110.0	20	2	89664	2.	97041600
250.0	T	32	13.	-49.3	0.0	-49.3	22	2	89664	2.	97041600

Fig. 16 Standard print output for a report with errors, containing quick-recognition part, main part, and list of events that will be included in prepbufr file.

9. Examples of operation of CQCHT96

Several examples are given to illustrate the operation of cqcht96, first showing elementary corrections, then more complicated corrections and observation errors. The first example, Fig. 17, shows a correction to a single height. The values for the observation and each check are separated. For each check, the value for height is first, followed by the value for temperature. The value of the hydrostatic residual applies between the level at which the value is placed and the next (complete) mandatory level above. Highlighted values indicate large residuals. Note the negative influence of the bad height value upon the vertical checks at 150 and 70 hPa. The correction in this case is -600, which is simple. The values after the correction are shown below and the highlighted values are now small.

CORRECTION TO SINGLE HEIGHT
Stn: 40754 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
200	11790	-41.3	2		2	3.4	-11	2.4	7	1.3
150	13690	-54.3	602		4	-1.4	-6	-1.5	-191	-1.6
119		-61.9				-1.8				-1.6
100	16820	-63.7	-600		593	0.3	581		590	1.5
90.8		-68.1				-2.1				-2.9
71.3		-62.5				3.6				2.4
70	18390	-63.5	-8		4	2.5			-209	0.8

↓

100 hPa height correction: -600m

↓

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
200	11790	-41.3	2		2	3.4	-11	2.4	7	1.3
150	13690	-54.3	2		4	-1.4	-6	-1.5	5	-1.6
119		-61.9				-1.8				-1.6
100	16220	-63.7	0		-7	0.3	-19		-10	1.5
90.8		-68.1				-2.1				-2.9
71.3		-62.5				3.6				2.4
70	18390	-63.5	-8		4	2.5			7	0.8

Fig. 17 Correction to a single height at station 40754 for 00 UTC 6 February 1997. A simple correction of -600 m is applied.

The next example, shown in Fig. 18, shows a correction to a single temperature at 200 hPa. The hydrostatic residuals (Xs, actually) indicate an error of exactly 31.5 degrees. The simple correction near this value, -30.0, is applied. The corrected profile and the new residual values is shown below.

CORRECTION TO SINGLE TEMPERATURE
Stn: 96163 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
300	9590	-35.3		1.8	-52	-1.6	-30	-1.1	-22	-1.1
250	10840	-44.3		-31.5	-52	0.1	-23	0.2	-9	-6.3
200	12290	-26.5		-31.5	-60	28.9	-29	29.2	-14	28.9
150	14060	-67.9		-0.2	-70	-0.3	-27	0.7	-27	6.8

200 hPa temperature correction: -30.0 deg

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
300	9590	-35.3		1.8	-52	-1.6	-30	-1.1	-22	-1.1
250	10840	-44.3		-1.5	-52	0.1	-23	0.2	-9	0.8
200	12290	-56.5		-1.5	-60	-1.1	-29	-0.8	-14	-1.1
150	14060	-67.9		-0.2	-70	-0.3	-27	0.7	-27	0.1

Fig. 18 Single temperature correction at 200 hPa for station 96163 at 00 UTC 6 February 1997.

Much more complicated corrections are possible with cqcht96. Fig. 19 shows such a case. Careful examination indicates that there were two errors made at 700 hPa: a temperature communication error of about -40. degrees and a computation error of 200-250 m between 850 and 700 hPa. The corrections that were made fit this pattern, but they were actually made in two steps. When the examination for errors reached 700 hPa, the height and temperature at this level were corrected (by 220 meters and 40.0 degrees). At this point, the only remaining error would appear to be a computation error between 700 and 500 hPa, and just such an apparent error was corrected next. The residuals after the corrections are all small.

TEMPERATURE AND HEIGHT CORRECTIONS TO SINGLE LEVEL COMPUTATION ERROR CORRECTION

Stn: 52495 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
870	1329	-14.7			3	-4.9			6	-4.3
850	1502	-8.5	-110	38.5	-4	-0.9	12	-0.4	67	8.0
700	2763	-55.5	215	43.7	-240	-41.2	-227	-39.7	-164	-41.0
500	5250	-29.5	-17	-2.2	-230	0.5	-218	-0.1	-66	8.6
300	8700	-53.1	-4	-1.4	-219	0.7	-211	1.8	-55	0.5
250	9860	-57.1	-1	-2.2	-216	0.3	-202	2.0	-47	0.7
234		-59.3				-1.2				-1.3
200	11260	-58.3	10	2.3	-218	-0.1	-198	-0.4	-62	-0.1
150	13100	-53.3	4	0.7	-194	2.2	-189	1.7	-50	2.2
100	15710	-53.9			-170	0.2	-179	-1.1	-76	-0.2

↓

700 hPa height correction: 220m
700 hPa temperature correction: 40.0 deg
500-100 hPa height corrections: 200m

↓

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
870	1329	-14.7			3	-4.9			6	-4.3
850	1502	-8.5	-4	-1.1	-4	-0.9	12	-0.4	67	2.1
700	2983	-15.5	-2	-0.3	-20	-1.2	-7	0.3	-164	-1.0
500	5450	-29.5	-17	-2.2	-30	0.5	-18	-0.1	-66	0.6
300	8900	-53.1	-4	-1.4	-19	0.7	-11	1.8	-55	0.5
250	10060	-57.1	-1	-2.2	-16	0.3	-2	2.0	-47	0.7
234		-59.3				-1.2				-1.3
200	11460	-58.3	10	2.3	-18	-0.1	2	-0.4	-62	-0.1
150	13300	-53.3	4	0.7	6	2.2	11	1.7	-50	2.2
100	15910	-53.9			30	0.2	21	-1.1	-76	-0.2

Fig. 19 Compound corrections: temperature communication and computation errors at the same level. Data from station 52495 for 00 UTC 6 February 1997.

Fig. 20 shows an example of a correction at the top mandatory level. At such a level, only the hydrostatic residual below is available, so that the necessary correction(s) is(are) ambiguous. There may be a correction to height, temperature or both required. Other residuals must be used to determine what to correct. In this case, the temperature is corrected by 20.0 degrees. The profile and the resultant residuals is shown below. All residuals become acceptable.

CORRECTION TO TOP LEVEL TEMPERATURE
Stn: 42667 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**
250	10810	-42.1	66	20.2	-71	-1.5	-29	-1.7	-15	-0.7
216		-49.1				-1.3				8.5
200	12280	-74.3			-84	-22.6	-41	-21.2	-44	-20.4
187		-59.9				-4.6				6.9

↓

200 hPa temperature correction: 20.0 deg

↓

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**
250	10810	-42.1	66	0.2	-71	-1.5	-29	-1.7	-15	-0.7
216		-49.1				-1.3				0.2
200	12280	-54.3			-84	-2.6	-41	-1.2	-44	-0.4
187		-59.9				-4.6				-3.3

Fig. 20 Correction to the top mandatory level data—temperature. The data are from station 42667 for 00 UTC 6 February 1997.

A more complicated case is shown in Fig. 21 where 3 communication errors and a computation error, affecting 2 levels, are corrected. The figure only shows the necessary parts of the profile. At 925 hPa there is a height error. Since it is separated vertically from the other errors, it is easily solved as an isolated height error. There are also temperature errors at 288 and 176 hPa. The temperature error at 228 hPa is diagnosed to be a communication error, and therefore correctable, because of the large hydrostatic residual for the 300-250 hPa layer. The temperature at 176 hPa is similarly corrected, but only after the computation error between 250 and 200 hPa is corrected and a large hydrostatic residual remains. Note that all the corrections are simple: 750 to 790, 11.3 to -41.3, 11790 to 11990 (and 13640 to 13840), and 51.4 to -51.4. The corrections are not always simple, but the majority are.

MULTIPLE CORRECTIONS
Stn: 42314 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
1003	111	14.8			-40	-0.2			-15	0.0
1000	138	14.6	-45		-38	-0.3			8	-0.3
925	750	11.4	44		-79	0.7			-52	0.4
924		11.4				0.7				0.4
850	1498	8.4	4		-30	-0.6			16	-0.5
...	...	...	...	...	...	...	...	...	...	...
300	9270	-41.7	-148		-71	-2.9			-24	-28.0
288		11.3				50.9				52.2
250	10500	-41.5	-204		-82	0.9			51	-14.9
212		-47.5				-0.7				-28.5
200	11790	-40.3	-448		-284				-151	
176		51.4				104.6				105.3
150	13640	-61.1			-283	-1.8			-131	-38.8

925 hPa height correction: 40m (communication)
288 hPa temperature correction: -52.6 deg
200 hPa height correction: 200m (computation)
150 hPa height correction: 200m (computation)
176 hPa temperature correction: 102.8 deg

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
1003	111	14.8			-40	-0.2			-15	0.0
1000	138	14.6	-5		-38	-0.3			8	-0.3
925	790	11.4	4		-39	0.7			-12	0.4
924		11.4				0.7				0.4
850	1498	8.4	4		-30	-0.6			16	-0.5
...	...	...	...	...	...	...	...	...	...	...
300	9270	-41.7	-7		-71	-2.9			-24	-1.3
288		-41.3				-1.7				-0.4
250	10500	-41.5	-4		-82	0.9			51	1.6
212		-47.5				-0.7				-1.4
200	11990	-40.3	-15		-84				-151	
176		-51.4				1.8				2.5
150	13840	-61.1			-83	-1.8			-131	-2.4

Fig. 21 Multiple corrections at a single station, 42314 at 00 UTC 6 February 1997. The corrections are for two temperature errors, a height communication error and a computation error.

There are several different possible baseline errors. Fig. 22 shows the simplest of these to correct, a surface pressure communication error correction. The surface pressure increment (PIS) is large and agrees with the baseline residual in terms of pressure (PSINC). Likewise, the baseline residual (BASRES) is large and agrees, with opposite sign, with the increment of height at the surface (965 hPa). A correction of 20.0 hPa is made and the lower part of the figure shows the resulting data and residual values.

SURFACE PRESSURE COMMUNICATION ERROR CORRECTION
Stn: 61052 Date: 97022800

INITIAL PROFILE
PIS: -17.7 PSINC: -18.7 BASRES: 169.

	observed		hyd resid		increment		horizontal		vertical	
pres	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**
1000	65				-12				94	
965	233	27.8			-162	-4.1			-160	-3.1
925	778	28.3	-7		6	-1.8			87	0.4

↓

surface pressure correction: 20.0 hPa

↓

CORRECTED PROFILE
PIS: 2.3 PSINC: 1.4 BASRES: -13.

	corrected		hyd resid		increment		horizontal		vertical	
pres	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**
1000	65				-12				94	
985	233	27.8			21	-4.1			23	-3.1
925	778	28.3	-7		6	-1.8			87	0.4

Fig. 22 Surface pressure communication error correction for station 61052 at 00 UTC 28 February 1997.

The final example, in Fig. 23, shows the most usual type of observation errors. Throughout most of the profile there is a fairly small negative temperature increment. When accumulated over a considerable depth of the atmosphere, this leads to large height increments. In cases like this, the question is which levels of height and which levels of temperature to flag. The way that cqcht96 is presently configured, it flagged the 150-20 hPa heights only. This is certainly not correct since the heights cannot be bad except through the influence of bad temperatures. But which should be flagged?

OBSERVATION ERRORS, SEVERAL HEIGHTS
Stn: 43371 Date: 97022800

DATA PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**
1004	64	23.8			-1	-3.9			-2	-1.4
1000	102	23.6	1		2	-3.8			2	-1.2
925	781	18.8	0		2	-2.2			0	-0.3
902		17.6				-1.7				-0.2
874		16.4				-1.5				-0.3
850	1504	16.0	0		3	-1.3			4	-0.5
803		14.3				-0.4				1.3
705		4.6				-4.4				-2.0
700	3121	4.8	-3		-5	-4.0			3	-1.3
678		5.8				-2.0				0.5
600		1.3				-2.0				-0.5
584		0.0				-2.1				-0.6
515		-5.8				-2.3				-0.2
500	5810	-8.1	4		-30	-3.2			-8	-1.6
400	7510	-18.9	-3		-49	-2.5			-12	-1.2
345		-26.9				-1.8				-0.2
300	9580	-35.7	1		-71	-2.7			-17	-1.0
250	10820	-46.3	1		-88	-4.2			-15	-1.8
211		-56.5				-5.7				-2.5
200	12260	-58.3	-4		-121	-4.8			-27	-1.4
150	14010	-71.4	6		-164	-3.5			-50	-1.3
112		-83.2				-5.5				-2.9
100	16320	-83.9	1		-198	-5.1			-46	-1.4
78.6		-83.2				-8.0				-5.0
70	18310	-78.2	-2		-254	-5.3			-78	-1.6
68.6		-76.5	8			-3.8				-0.4
50	20280	-68.7	5		-281	-1.8			-92	-1.0
30	23420	-58.7			-281	-0.1			-98	-1.3
26.5		-52.9				3.4				3.6
20	26040	-50.1			-245	-0.7			-108	-1.6
17.9		-47.9				0.7				1.1
13.9		-49.3				-0.9				-1.1

↓

150-70 hPa heights questionable
70 hPa height bad
30-20 hPa heights questionable

Fig. 23 Observation errors at several levels for station 43371 at 00 UTC 28 February 1997

In other examples, the temperature increments are large, with possibly reversing sign, only over a range of pressures and good elsewhere. This may lead to heights that likewise have large increments only over a limited pressure range. It seems reasonable, however, that if bad temperatures at any level lead to a bad height, then all heights above should also be flagged. While the present version of cqcht96 does not operate this way, perhaps it should in the future.

10. Performance of CQCHT96 for January 1997

Previous monitoring efforts of rawinsonde quality were reported in Morone, et al, 1992. The result reported here apply to January 1997 when there were an average of 587 land stations and 9 ships with rawinsonde reports available at OO and 12 UTC to NCEP. At 00 UTC there were somewhat more reports (616 land and 8 ship), compared to 12 UTC (558 land and 9 ship).

Table 2. Percent Distribution of Error Suspicions for Land Stations, Percent of Land Stations Having One or More Error Suspicions and Average Number of Land Stations with One or More Error Suspicions by Type

Error Type	Percent of total	Percent of stns	Number
height at one level	3.4	2.0	11.5
temperature at one level	1.4	.8	4.8
height and temperature	1.8	1.0	6.1
top or lower hole boundary	1.7	1.0	5.7
computation error in height	5.7	3.3	19.2
2 heights	.2	.1	.7
2 temperatures	.2	.1	.6
lower height, upper temperature	.1	.0	.2
lower temperature, upper height	.1	.0	.2
significant level temperature corr.	4.2	2.4	14.2
sig. level temperature, not corrected	1.9	1.1	6.4
observation error (T or z)	77.8	44.6	261.9
surface press. Communication error	.7	.4	2.4
surface temperature error	.2	.1	.6
surface undetermined error	.4	.3	1.5
surface press. observation error	.1	.0	.2
TOTAL	**100.**	**(57.2)**	**(336.2)**

Table 2 shows the distribution of errors among the various types. The first column, percent of total, gives the distribution among the error types, and sums to 100%. The next column tells the percent of stations containing, on average, a particular type of error. For most error types, there are only a few percent of the stations that have them. However, for observation errors, about 45% of stations have at least 1. These numbers cannot strictly be added since a station may have more than one kind of error. The last column converts the percent of stations with a particular error type into a number of stations. This number is just the % times 587.

Acknowledgments

I would like to acknowledge that this work is the result of the pioneering work of Lev Gandin in the Soviet Union and his later emigration to the United States and desire to continue his work here on quality control. Since his arrival, we have constantly worked together on a framework already well established in his mind, with much of the detail resulting from true collaboration. It has been a privilege; thank you.

References

Collins, W.G. and L.S. Gandin, 1990: Comprehensive quality control at the National Meteorological Center *MWR*, **118**,2752-2767.

Collins, W.G. and L.S. Gandin, 1992: Complex quality control of rawinsonde heights and temperatures (CQCHT) at the National Meteorological Center, Office Note 390, September 1992.

Gandin, L.S., 1988: Complex quality control of meteorological observations, *MWR*, **116**, 1137-1156.

Gandin, L.S., W.G. Collins and L.L. Morone, 1993: Rough errors in rawinsonde reports: present situation and ways to improve it. Preprint Volume, *13th Conference on Weather Analysis and Forecasting*, 2-6 August, 1993, Vienna, VA, 254-255.

Morone, L.L., L.S. Gandin and W.G. Collins, 1992: Quasi-operational monitoring of the NMC complex quality control performance. Preprints, Twelfth Conference on Probability and Statistics in the Atmospheric Sciences, American Meteorological Society, Toronto, Ont., Canada, 22-26 June, 1992.

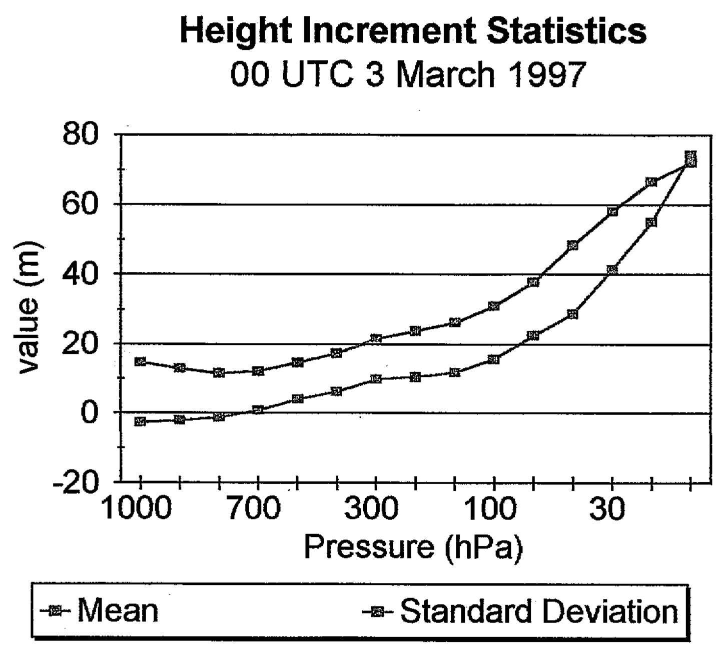

Fig. 3 Height increment mean and standard deviation for all stations for a single time: 00 UTC 3 March 1997.

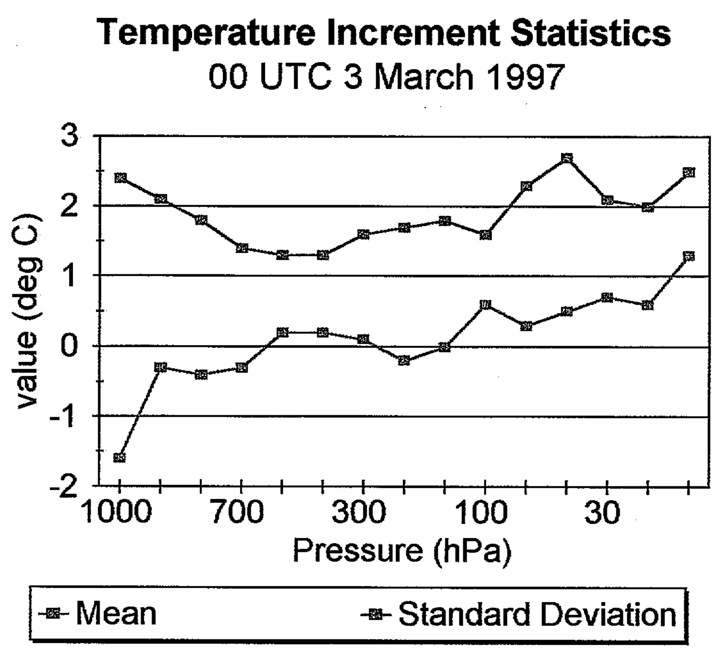

Fig. 4 Temperature increment mean and standard deviation for all stations for a single time: 00 UTC 3 March 1997.

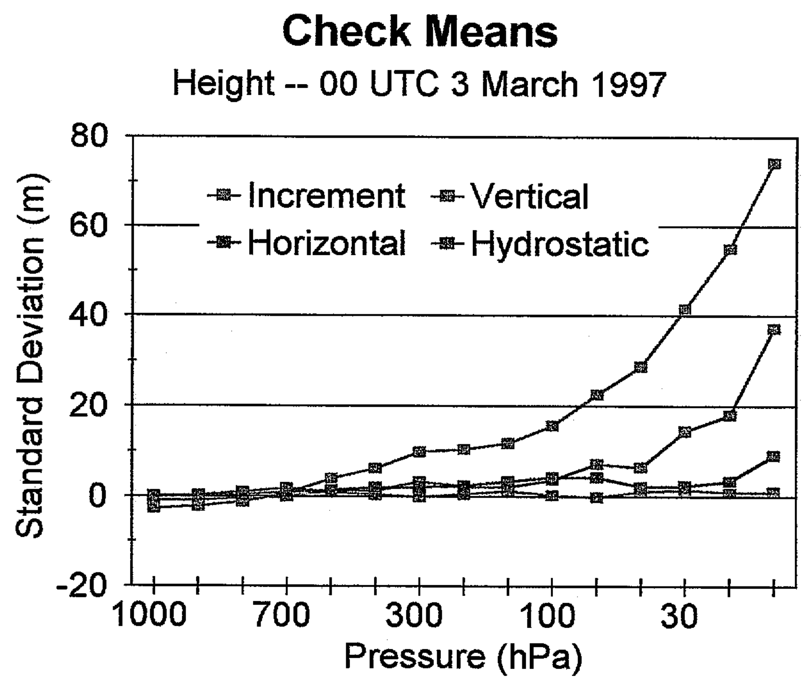

Fig. 11 Check means for height for a sample time: 00UTC 3 March 1997. Includes increment, horizontal, vertical and hydrostatic checks.

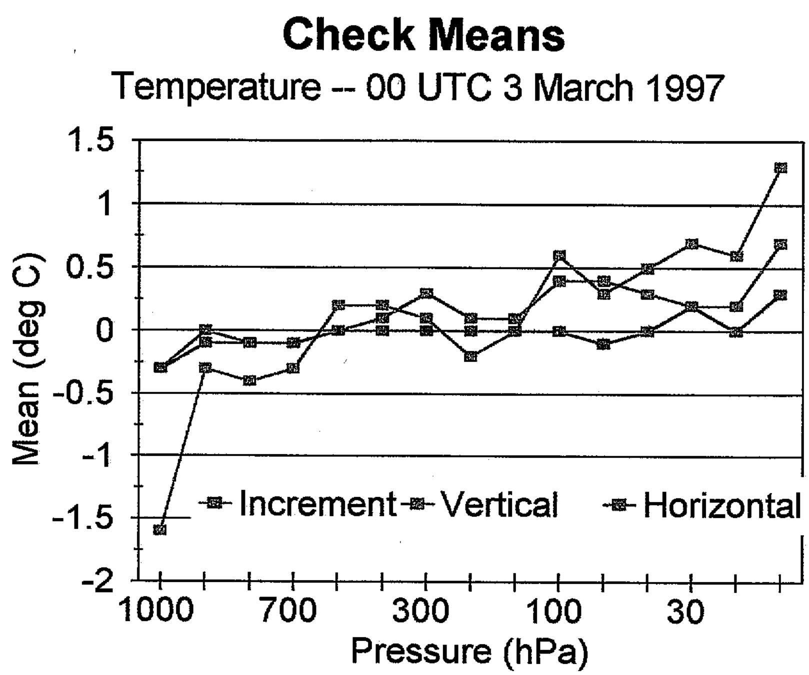

Fig. 12 Check means for temperature for a sample time: 00UTC 3 March 1997.

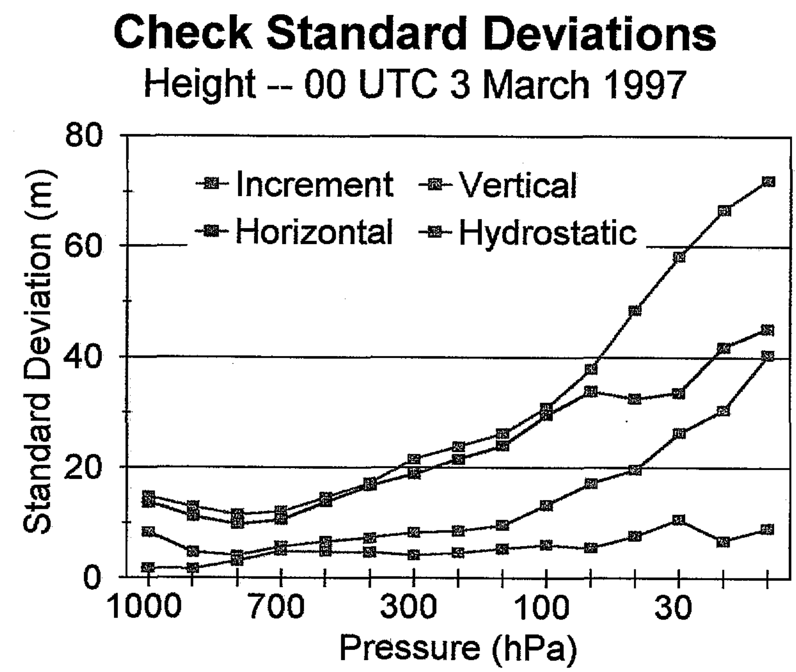

Fig. 13 Height check standard deviations for a sample time: 00UTC 3 March 1997.

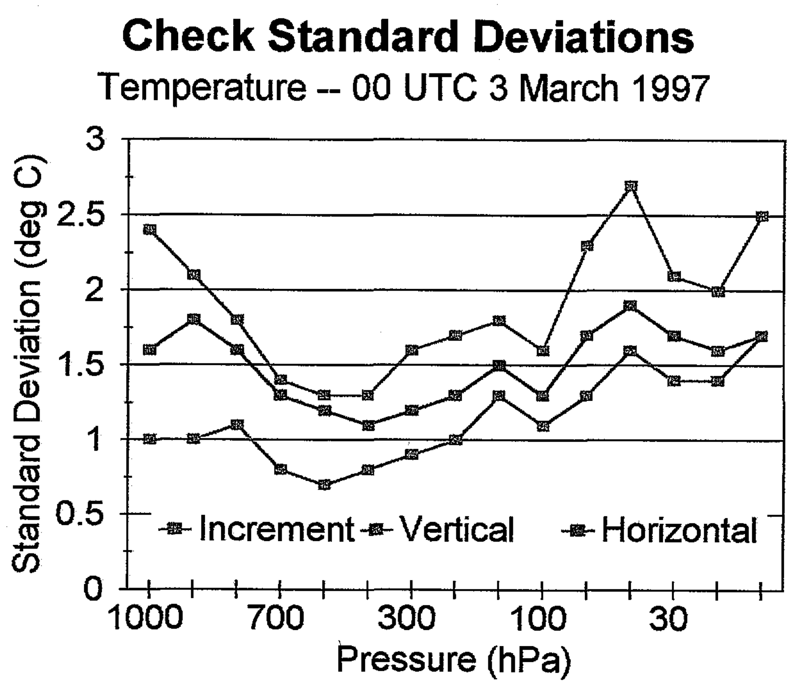

Fig. 13 Temperature check standard deviations for a sample time: 00UTC 3 March 1997.

next (complete) mandatory level above. Highlighted values indicate large residuals. Note the negative influence of the bad height value upon the vertical checks at 150 and 70 hPa. The correction in this case is -600, which is simple. The values after the correction are shown below and the highlighted values are now small.

CORRECTION TO SINGLE HEIGHT
Stn: 40754 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
200	11790	-41.3	2		2	3.4	-11	2.4	7	1.3
150	13690	-54.3	602		4	-1.4	-6	-1.5	191	-1.6
119		-61.9				-1.8				-1.6
100	16820	-63.7	-600		593	0.3	581		590	1.5
90.8		-68.1				-2.1				-2.9
71.3		-62.5				3.6				2.4
70	18390	-63.5	-8		4	2.5			209	0.8

↓

100 hPa height correction: -600m

↓

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
200	11790	-41.3	2		2	3.4	-11	2.4	7	1.3
150	13690	-54.3	2		4	-1.4	-6	-1.5	5	-1.6
119		-61.9				-1.8				-1.6
100	16220	-63.7	0		-7	0.3	-19		-10	1.5
90.8		-68.1				-2.1				-2.9
71.3		-62.5				3.6				2.4
70	18390	-63.5	-8		4	2.5			7	0.8

Fig. 17 Correction to a single height at station 40754 for 00 UTC 6 February 1997. A simple correction of -600 m is applied.

The next example, shown in Fig. 18, shows a correction to a single temperature at 200 hPa. The hydrostatic residuals (Xs, actually) indicate an error of exactly 31.5 degrees. The simple correction near this value, -30.0, is applied. The corrected profile and the new residual values is shown below.

CORRECTION TO SINGLE TEMPERATURE
Stn: 96163 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**
300	9590	-35.3		1.8	-52	-1.6	-30	-1.1	-22	-1.1
250	10840	-44.3		-31.5	-52	0.1	-23	0.2	-9	6.3
200	12290	-26.5		-31.5	-60	28.9	-29	29.2	-14	28.9
150	14060	-67.9		-0.2	-70	-0.3	-27	0.7	-27	6.8

↓

200 hPa temperature correction: -30.0 deg

↓

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**
300	9590	-35.3		1.8	-52	-1.6	-30	-1.1	-22	-1.1
250	10840	-44.3		-1.5	-52	0.1	-23	0.2	-9	0.8
200	12290	-56.5		-1.5	-60	-1.1	-29	-0.8	-14	**-1.1**
150	14060	-67.9		-0.2	-70	-0.3	-27	0.7	-27	0.1

Fig. 18 Single temperature correction at 200 hPa for station 96163 at 00 UTC 6 February 1997.

Much more complicated corrections are possible with cqcht96. Fig. 19 shows such a case. Careful examination indicates that there were two errors made at 700 hPa: a temperature communication error of about -40. degrees and a computation error of 200-250 m between 850 and 700 hPa. The corrections that were made fit this pattern, but they were actually made in two steps. When the examination for errors reached 700 hPa, the height and temperature at this level were corrected (by 220 meters and 40.0 degrees). At this point, the only remaining error would appear to be a computation error between 700 and 500 hPa, and just such an apparent error was corrected next. The residuals after the corrections are all small.

TEMPERATURE AND HEIGHT CORRECTIONS TO SINGLE LEVEL COMPUTATION ERROR CORRECTION
Stn: 52495 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**
870	1329	-14.7			3	-4.9			6	-4.3
850	1502	-8.5	-110	-38.5	-4	-0.9	12	-0.4	67	8.0
700	2763	-55.5	215	43.7	-240	-41.2	-227	-39.7	-164	-41.0
500	5250	-29.5	-17	-2.2	-230	0.5	-218	-0.1	-66	8.6
300	8700	-53.1	-4	-1.4	-219	0.7	-211	1.8	-55	0.5
250	9860	-57.1	-1	-2.2	-216	0.3	-202	2.0	-47	0.7
234		-59.3				-1.2				-1.3
200	11260	-58.3	10	2.3	-218	-0.1	-198	-0.4	-62	-0.1
150	13100	-53.3	4	0.7	-194	2.2	-189	1.7	-50	2.2
100	15710	-53.9			-170	0.2	-179	-1.1	-76	-0.2

↓

700 hPa height correction: 220m
700 hPa temperature correction: 40.0 deg
500-100 hPa height corrections: 200m

↓

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**	**z**	**T**
870	1329	-14.7			3	-4.9			6	-4.3
850	1502	-8.5	-4	-1.1	-4	-0.9	12	-0.4	67	2.1
700	2983	-15.5	-2	-0.3	-20	-1.2	-7	0.3	-164	-1.0
500	5450	-29.5	-17	-2.2	-30	0.5	-18	-0.1	-66	0.6
300	8900	-53.1	-4	-1.4	-19	0.7	-11	1.8	-55	0.5
250	10060	-57.1	-1	-2.2	-16	0.3	-2	2.0	-47	0.7
234		-59.3				-1.2				-1.3
200	11460	-58.3	10	2.3	-18	-0.1	2	-0.4	-62	-0.1
150	13300	-53.3	4	0.7	6	2.2	11	1.7	-50	2.2
100	15910	-53.9			30	0.2	21	-1.1	-76	-0.2

Fig. 19 Compound corrections: temperature communication and computation errors at the same level. Data from station 52495 for 00 UTC 6 February 1997.

Fig. 20 shows an example of a correction at the top mandatory level. At such a level, only the hydrostatic residual below is available, so that the necessary correction(s) is(are) ambiguous. There may be a correction to height, temperature or both required. Other residuals must be used to determine what to correct. In this case, the temperature is corrected by 20.0 degrees. The profile and the resultant residuals is shown below. All residuals become acceptable.

CORRECTION TO TOP LEVEL TEMPERATURE
Stn: 42667 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
250	10810	-42.1	66	20.2	-71	-1.5	-29	-1.7	-15	-0.7
216		-49.1				-1.3				8.5
200	12280	-74.3			-84	-22.6	-41	-21.2	-44	-20.4
187		-59.9				-4.6				6.9

↓

200 hPa temperature correction: 20.0 deg

↓

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
250	10810	-42.1	66	0.2	-71	-1.5	-29	-1.7	-15	-0.7
216		-49.1				-1.3				0.2
200	12280	-54.3			-84	-2.6	-41	-1.2	-44	-0.4
187		-59.9				-4.6				-3.3

Fig. 20 Correction to the top mandatory level data—temperature. The data are from station 42667 for 00 UTC 6 February 1997.

A more complicated case is shown in Fig. 21 where 3 communication errors and a computation error, affecting 2 levels, are corrected. The figure only shows the necessary parts of the profile. At 925 hPa there is a height error. Since it is separated vertically from the other errors, it is easily solved as an isolated height error. There are also temperature errors at 288 and 176 hPa. The temperature error at 228 hPa is diagnosed to be a communication error, and therefore correctable, because of the large hydrostatic residual for the 300-250 hPa layer. The temperature at 176 hPa is similarly corrected, but only after the computation error between 250 and 200 hPa is corrected and a large hydrostatic residual remains. Note that all the corrections are simple: 750 to 790, 11.3 to -41.3, 11790 to 11990 (and 13640 to 13840), and 51.4 to -51.4. The corrections are not always simple, but the majority are.

MULTIPLE CORRECTIONS
Stn: 42314 Date: 97020600

INITIAL PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
1003	111	14.8			-40	-0.2			-15	0.0
1000	138	14.6	-45		-38	-0.3			8	-0.3
925	750	11.4	44		-79	0.7			-52	0.4
924		11.4				0.7				0.4
850	1498	8.4	4		-30	-0.6			16	-0.5
...	...	...	...	...	...	...	...	...	...	...
300	9270	-41.7	-148		-71	-2.9			-24	-28.0
288		11.3				50.9				52.2
250	10500	-41.5	204		-82	0.9			51	-14.9
212		-47.5				-0.7				-28.5
200	11790	-40.3	-448		-284				-151	
176		51.4				104.6				105.3
150	13640	-61.1			-283	-1.8			-131	-38.8

↓

925 hPa height correction: 40m (communication)
288 hPa temperature correction: -52.6 deg
200 hPa height correction: 200m (computation)
150 hPa height correction: 200m (computation)
176 hPa temperature correction: 102.8 deg

↓

CORRECTED PROFILE

	corrected		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
1003	111	14.8			-40	-0.2			-15	0.0
1000	138	14.6	-5		-38	-0.3			8	-0.3
925	790	11.4	4		-39	0.7			-12	0.4
924		11.4				0.7				0.4
850	1498	8.4	4		-30	-0.6			16	-0.5
...	...	...	...	...	...	...	...	...	...	...
300	9270	-41.7	-7		-71	-2.9			-24	-1.3
288		-41.3				-1.7				-0.4
250	10500	-41.5	-4		-82	0.9			51	1.6
212		-47.5				-0.7				-1.4
200	11990	-40.3	-15		-84				-151	
176		-51.4				1.8				2.5
150	13840	-61.1			-83	-1.8			-131	-2.4

Fig. 21 Multiple corrections at a single station, 42314 at 00 UTC 6 February 1997. The corrections are for two temperature errors, a height communication error and a computation error.

There are several different possible baseline errors. Fig. 22 shows the simplest of these to correct, a surface pressure communication error correction. The surface pressure increment (PIS) is large and agrees with the baseline residual in terms of pressure (PSINC). Likewise, the baseline residual (BASRES) is large and agrees, with opposite sign, with the increment of height at the surface (965 hPa). A correction of 20.0 hPa is made and the lower part of the figure shows the resulting data and residual values.

SURFACE PRESSURE COMMUNICATION ERROR CORRECTION
Stn: 61052 Date: 97022800

INITIAL PROFILE
PIS: -17.7 PSINC: -18.7 BASRES: 169.

	observed		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
1000	65				-12				94	
965	233	27.8			-162	-4.1			-160	-3.1
925	778	28.3	-7		6	-1.8			87	0.4

↓

surface pressure correction: 20.0 hPa

↓

CORRECTED PROFILE
PIS: 2.3 PSINC: 1.4 BASRES: -13.

	corrected		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
1000	65				-12				94	
985	233	27.8			21	-4.1			23	-3.1
925	778	28.3	-7		6	-1.8			87	0.4

Fig. 22 Surface pressure communication error correction for station 61052 at 00 UTC 28 February 1997.

The final example, in Fig. 23, shows the most usual type of observation errors. Throughout most of the profile there is a fairly small negative temperature increment. When accumulated over a considerable depth of the atmosphere, this leads to large height increments. In cases like this, the question is which levels of height and which levels of temperature to flag. The way that cqcht96 is presently configured, it flagged the 150-20 hPa heights only. This is certainly not correct since the heights cannot be bad except through the influence of bad temperatures. But which should be flagged?

OBSERVATION ERRORS, SEVERAL HEIGHTS
Stn: 43371 Date: 97022800

DATA PROFILE

	observed		hyd resid		increment		horizontal		vertical	
pres	z	T	z	T	z	T	z	T	z	T
1004	64	23.8			-1	-3.9			-2	-1.4
1000	102	23.6	1		2	-3.8			2	-1.2
925	781	18.8	0		2	-2.2			0	-0.3
902		17.6				-1.7				-0.2
874		16.4				-1.5				-0.3
850	1504	16.0	0		3	-1.3			4	-0.5
803		14.3				-0.4				1.3
705		4.6				-4.4				-2.0
700	3121	4.8	-3		-5	-4.0			3	-1.3
678		5.8				-2.0				0.5
600		1.3				-2.0				-0.5
584		0.0				-2.1				-0.6
515		-5.8				-2.3				-0.2
500	5810	-8.1	4		-30	-3.2			-8	-1.6
400	7510	-18.9	-3		-49	-2.5			-12	-1.2
345		-26.9				-1.8				-0.2
300	9580	-35.7	1		-71	-2.7			-17	-1.0
250	10820	-46.3	1		-88	-4.2			-15	-1.8
211		-56.5				-5.7				-2.5
200	12260	-58.3	-4		-121	-4.8			-27	-1.4
150	14010	-71.4	6		-164	-3.5			-50	-1.3
112		-83.2				-5.5				-2.9
100	16320	-83.9	1		-198	-5.1			-46	-1.4
78.6		-83.2				-8.0				-5.0
70	18310	-78.2	-2		-254	-5.3			-78	-1.6
68.6		-76.5	8			-3.8				-0.4
50	20280	-68.7	5		-281	-1.8			-92	-1.0
30	23420	-58.7			-281	-0.1			-98	-1.3
26.5		-52.9				3.4				3.6
20	26040	-50.1			-245	-0.7			-108	-1.6
17.9		-47.9				0.7				1.1
13.9		-49.3				-0.9				-1.1

↓

150-70 hPa heights questionable
70 hPa height bad
30-20 hPa heights questionable

Fig. 23 Observation errors at several levels for station 43371 at 00 UTC 28 February 1997

In other examples, the temperature increments are large, with possibly reversing sign, only over a range of pressures and good elsewhere. This may lead to heights that likewise have large increments only over a limited pressure range. It seems reasonable, however, that if bad temperatures at any level lead to a bad height, then all heights above should also be flagged. While the present version of cqcht96 does not operate this way, perhaps it should in the future.

10. Performance of CQCHT96 for January 1997

Previous monitoring efforts of rawinsonde quality were reported in Morone, et al, 1992. The result reported here apply to January 1997 when there were an average of 587 land stations and 9 ships with rawinsonde reports available at OO and 12 UTC to NCEP. At 00 UTC there were somewhat more reports (616 land and 8 ship), compared to 12 UTC (558 land and 9 ship).

Table 2. Percent Distribution of Error Suspicions for Land Stations, Percent of Land Stations Having One or More Error Suspicions and Average Number of Land Stations with One or More Error Suspicions by Type

Error Type	Percent of total	Percent of stns	Number
height at one level	3.4	2.0	11.5
temperature at one level	1.4	.8	4.8
height and temperature	1.8	1.0	6.1
top or lower hole boundary	1.7	1.0	5.7
computation error in height	5.7	3.3	19.2
2 heights	.2	.1	.7
2 temperatures	.2	.1	.6
lower height, upper temperature	.1	.0	.2
lower temperature, upper height	.1	.0	.2
significant level temperature corr.	4.2	2.4	14.2
sig. level temperature, not corrected	1.9	1.1	6.4
observation error (T or z)	77.8	44.6	261.9
surface press. Communication error	.7	.4	2.4
surface temperature error	.2	.1	.6
surface undetermined error	.4	.3	1.5
surface press. observation error	.1	.0	.2
TOTAL	100.	(57.2)	(336.2)

Table 2 shows the distribution of errors among the various types. The first column, percent of total, gives the distribution among the error types, and sums to 100%. The next column tells the percent of stations containing, on average, a particular type of error. For most error types, there are only a few percent of the stations that have them. However, for observation errors, about 45% of stations have at least 1. These numbers cannot strictly be added since a station may have more than one kind of error. The last column converts the percent of stations with a particular error type into a number of stations. This number is just the % times 587.

Acknowledgments

I would like to acknowledge that this work is the result of the pioneering work of Lev Gandin in the Soviet Union and his later emigration to the United States and desire to continue his work here on quality control. Since his arrival, we have constantly worked together on a framework already well established in his mind, with much of the detail resulting from true collaboration. It has been a privilege; thank you.

References

Collins, W.G. and L.S. Gandin, 1990: Comprehensive quality control at the National Meteorological Center *MWR*, **118**,2752-2767.

Collins, W.G. and L.S. Gandin, 1992: Complex quality control of rawinsonde heights and temperatures (CQCHT) at the National Meteorological Center, Office Note 390, September 1992.

Gandin, L.S., 1988: Complex quality control of meteorological observations, *MWR*, **116**, 1137-1156.

Gandin, L.S., W.G. Collins and L.L. Morone, 1993: Rough errors in rawinsonde reports: present situation and ways to improve it. Preprint Volume, *13th Conference on Weather Analysis and Forecasting*, 2-6 August, 1993, Vienna, VA, 254-255.

Morone, L.L., L.S. Gandin and W.G. Collins, 1992: Quasi-operational monitoring of the NMC complex quality control performance. Preprints, Twelfth Conference on Probability and Statistics in the Atmospheric Sciences, American Meteorological Society, Toronto, Ont., Canada, 22-26 June, 1992.